1-1-1981

A study of the evaluation efforts of innovative educational programs funded under Title IV Part C of the Elementary and Secondary Education Act in Massachusetts.

Mary Antoinette Seroski
University of Massachusetts Amherst

A STUDY OF THE EVALUATION EFFORTS OF INNOVATIVE EDUCATIONAL PROGRAMS FUNDED UNDER TITLE IV PART C OF THE ELEMENTARY AND SECONDARY EDUCATION ACT IN MASSACHUSETTS

A Dissertation Presented

By

MARY ANTOINETTE SEROSKI

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

DOCTOR OF EDUCATION

February 1981

School of Education

A STUDY OF THE EVALUATION EFFORTS OF INNOVATIVE EDUCATIONAL PROGRAMS FUNDED UNDER TITLE IV PART C OF THE ELEMENTARY AND SECONDARY EDUCATION ACT IN MASSACHUSETTS

A Dissertation Presented

By

MARY ANTOINETTE SEROSKI

Approved as to style and content by:

Dr. William C. Wolf, Jr.,
Chairperson of Committee

Dr. Hariharan Swaminathan, Member

Dr. Ronnie Janoff-Bulman, Member

Dr. Mario Fantini, Dean
School of Education

D E D I C A T I O N

TO MY

PARENTS

A C K N O W L E D G E M E N T S

I am grateful to many individuals for their support and assistance throughout the completion of both this study and my doctoral program.

As both Chairperson of my Dissertation Committee and my major advisor, Dr. William C. Wolf, Jr., has been a constant source of help and advice from my initial interview through my exit interview. For his ever available time and assistance, I am truly grateful. Sincere thanks are owed to Dr. Hariharan Swaminathan, for his advice on the implementation of this study. Dr. Swaminathan's great attention to detail helped immensely in the final stages of the dissertation process. I am grateful to Dr. Ronnie Janoff-Bulman for her support and kindness and particularly for her help editing the final draft.

To Dr. Ralph Anttonen, Director, Stayer Research and Learning Center, Millersville State College, a note of appreciation for the encouragement he continually provided in the undertaking of my doctoral studies.

Thanks are offered to Dr. John F. Reynolds, Massachusetts Department of Education, ESEA Title IV-C Coordinator, who provided the initial encouragement for my pursuit of this study.

The instruction and advice of Dr. G. Ernest Anderson, in the early stages of this study is deeply appreciated. I am grateful to Mary Beth Regan, my graduate colleague in the Research and Evaluation Methods Program, who was a member of the committee for my comprehensive examination. I am indebted to Bernie McDonald for the skillful and aesthetic typing of this dissertation.

My deepest appreciation is extended to Nicholas Terzis for his encouragement and assistance. His optimism for the successful completion of this endeavor helped keep the dream of success within sight.

I am grateful to my siblings, Rita and Rick, for their support throughout this endeavor. I would like to acknowledge my two cats, Natosha and Sassafrass, who playfully rearranged data, tables, and drafts numerous times, for their unending love and attention. I offer my deepest thanks to my parents for their continual love and understanding and for encouraging me to always follow my interests. This project could not have been completed without their support.

A B S T R A C T

A Study of the Evaluation Efforts of Innovative Educational Programs Funded Under Title IV Part C of the Elementary and Secondary Education Act in Massachusetts

(February 1980)

Mary Seroski, B.A., Point Park College
M.S., Millersville State College
Ed.D., University of Massachuestts

Directed by: Professor William C. Wolf, Jr.

The Elementary and Secondary Education Act of 1965 has provided federal funds to over 100 school systems in Massachusetts to find creative solutions to local problems. It is the intent of the legislation to provide "seed money" to a school system to test a new solution to a local problem, in order to identify successful ventures through validation by the Joint Dissemination Review Panel (JDRP), resulting finally in the national dissemination of the innovation. Unhappily, not one of the Massachusetts Title IV-C projects have been presented to the JDRP. What has sadly unfolded is that these federally-funded projects cannot measure-up to established validation procedures. Millions of dollars were invested in innovative education projects in Massachusetts. The return on this investment in the form of validated projects is not impressive.

It was the purpose of the study to investigate the Massachusetts Title IV-C program efforts. The three major objectives of this study were: (1) to systematically study the state of Massachusetts' Title IV-C operation in order to document evaluation policies and practices; (2) to identify problems, based upon data obtained via surveys and document analysis, which contribute to the program's unsatisfactory evaluation outcomes; and (3) to offer an evaluation plan which is tailored to resolve evaluation problems identified within the state's Title IV-C program.

The first study objective was accomplished by designing a descriptive study permitting data to be gathered from every available source. Six data collection procedures were utilized in order to study the state of Title IV-C in Massachusetts. These include the Massachusetts Title IV-C Assessment, Massachusetts Validation Process, Interviews, Evaluation Design Review, and Assessment of Massachusetts Evaluation Needs.

In order to carry out the second study objective, problems which contribute to the program's unsatisfactory evaluation outcomes were identified. Examination of the many categories of investigation enabled the investigator to document policies and practices, and to identify problems which contribute to non-validation. The

categories of investigation include project director, program objectives, evaluation design, instrumentation, statistics, validation, and evaluators.

An evaluation plan tailored to resolve the evaluation problems was developed as a result of the third part of the study. Guidelines were developed, which will convey to project directors and project evaluators that the evaluation designs for Title IV-C projects must be of such a quality that it would enhance a project's capability of being initially validated for diffusion by the Massachusetts Validation Process and subsequently validated for national dissemination by the Joint Dissemination Review Panel.

The success of the plan to date includes the funding of the Evaluation Improvement Project to implement a Summative Evaluation Management System for Title IV-C projects in Massachusetts. The directors of projects that received initial funding in 1979 and 1980 are receiving individualized training and technical assistance in the area of program evaluation that their predecessors did not. The adoption of this program is in part due to the support of the Massachusetts Department of Education Title IV-C Coordinator and staff, who took ownership of the concept.

TABLE OF CONTENTS

LIST OF TABLES

C H A P T E R I

INTRODUCTION

Problem Statement

When Elliot Richardson was the secretary of HEW, he compared education to the process of drilling for oil. An oil company will invest heavily in identifying and drilling new oil wells. When a gusher is found the precious liquid gold is distributed around the world. Similarly educators will invest heavily in resources to identify new educational practices. Once a new practice is identified, more resources will be invested in the research and development of the educational innovation. However, Mr. Richardson suggested that once an educator hits precious gold the well is immediately capped, declared a local triumph and not distributed or shared with others.

One is not surprised that educational oil wells (innovative education projects) are frequently capped when one considers that evaluative information about the student outcomes of the educational innovations are scarce, resulting in innovative educational programs that do not prove worthy, and so only look like they are

providing a great service to America's youth. In the absence of evaluative information, an innovation will not stand up against scrutiny.

When the United States Congress passed the Elementary and Secondary Education Act in 1965, Congress mandated that E.S.E.A. Title III innovative education projects be evaluated. Few realized most educators did not command sufficient expertise in evaluation and research to implement the law. Since relatively few practitioners were able to effectively evaluate their programs, most did not yield conclusive data.

When the states and later the federal government established procedures to verify the calibre of project outcomes, most project directors were unable to meet even the most basic verification procedures. The federal government established a panel, called the Joint Dissemination Review Panel (JKRP), to oversee what came to be known as "the validation process." Relatively few innovative projects have successfully "passed" the state or federal review processes.

During the lifetime of Title III in Massachusetts only two projects have had evaluation designs that were comprehensive enough to produce data worthy of presentation to the JDRP. Of the two only one, Project Adventure, received national validation.

There have been well over 100 projects funded under Title IV-C in Massachusetts. Not one Title IV-C project has been presented to the JDRP, and only two are half heartedly considering application to the JDRP currently.

What has unfolded in the period between 1965, when Congress called for the generation of evaluation data, and the present, when most federally-funded projects cannot measure-up to established validation procedures, is not encouraging. In this period, hundreds of millions of dollars have been invested in innovative education projects. The return on this investment in the form of validated projects is not impressive.

More needs to be learned about the evaluation process associated with innovative education projects. If researchers were able to identify break-down points in the evaluation process as it unfolds, the problems identified could be sujected to systematic study and conceivably corrected. Researchers are, in fact, initiating such studies.

The problem addressed in this dissertation stems from what has been taking place in the ESEA Title IV-C program within Massachusetts. As has been previously mentioned, none of the more than 100 funded IV-C projects has been validated. More needs to be learned about the

State's Title IV-C operation in order to relate program operations to the feeble evaluation outcomes documented.

This dissertation is aimed at the state of the art of evaluation within the ESEA Title IV-C program as it is administered within the state of Massachusetts. Evaluation problems have been identified within this program which hopefully can be meaningfully confronted and resolved.

Objectives

The purposes of this study were: to ascertain why E.S.E.A. Title IV-C projects, which are funded and expedited within the state of Massachusetts routinely fail to measure up to validation guidelines established by the state and federal governments. And, to offer remedies, based upon data obtained, that are likely to resolve identified deficiencies. Specifically, the researcher proposed to do the following:

1. To systematically study the state of Massachusetts' Title IV-C operation in order to document evaluation policies and practices.

2. To identify problems, based upon data obtained via surveys and document analysis, which contribute to the program's unsatisfactory evaluation outcomes.

3. To offer an evaluation plan which is tailored to resolve evaluation problems identified within the state's Title IV-C program.

This was a descriptive study of the evaluation efforts of Title IV-C projects in Massachusetts. Objective 1 examined the state of the art of educational evaluation for validation in Title IV-C projects in Massachusetts. Descriptive data, drawn from specific project records and from persons associated with the projects, were gathered, analyzed, and reported. In Objective 2 the evaluation problems which a Title IV-C innovative project encounters during its funding period were identified. The data that were gathered in Objective 1 were sorted, and related to what is known about viable evaluation methodology in order to identify problems within the Title IV-C program. In Objective 3, a developmental evaluation plan for innovative projects was designed. The specific problems that were identified in Objective 2 were analyzed and an evaluation plan was set forth relevant to the problems identified.

Definition of Terms

The following terms are defined as they are to be used in the study.

Educational Innovation. Educational innovation is an educational idea, practice, or object perceived as new by an individual (Rogers & Shoemaker, 1971).

> It matters little, so far as human behavior is concerned, whether or

> not an idea is "objectively" new as measured by the lapse of time since its first use or discovery. It is the perceived or subjective newness of the idea for the individual that determines his reaction to it. If the idea seems new to an individual it is an innovation. (p. 19)

The concept gained a new importance during the Johnson administration when educational innovation was established as law under Title III of the Elementary and Secondary Education Act, innovation became synonymous with E.S.E.A. Title III and later E.S.E.A. Title IV-C. For the purposes of this dissertation an innovative educational program is one in which the curriculum is perceived as new by the individuals serviced by the Local Education Agency to which the E.S.E.A. Title IV-C grant was awarded. In other words, an innovative educational program is an E.S.E.A. Title IV-C project.

Evaluation. The Elementary and Secondary Education Act of 1965 included a proviso requiring educators to be accountable for the federal monies they recieve and to file an evaluation report for each grant showing what effects had resulted from the expenditure of the federal funds (Worthen & Sanders, 1973, p. 6). Title IV-C projects are required, by law, to evaluate.

Formative Evaluation. Formative evaluation takes place during the development of a program. Formative evaluation provides the decision maker with information

during the course of program development and execution for possible mid-course corrections to help assure that the program objectives are eventually met (CSDE, 1977, p. A6).

Summative Evaluation. Summative evaluation is an assessment of the extent to which a program causes changes in the desired direction in the target population (Rossi, 1979). For the purpose of this dissertation, summative evaluation shall be defined as a determination of the impact a Title IV-C project has on the children participating in the program. A summative evaluation is one that can reasonably prove the children participating in the innovation learned more as a result of that participation. The evaluation will produce the evidence necessary to determine the worthiness of the project for validation.

Validation. A quality control mechanism in the form of a review process.

Massachusetts Validation Process. The Massachusetts validation process is a process which determines the worth of a project by an on-site visit of one day by a validation team. When a project is validated for diffusion in Massachusetts, the Massachusetts Department of Education will diffuse the project throughout the state and may support an attempt for validation through the Joint Dissemination Review Panel (JDRP).

Joint Dissemination Review Panel. The JDRP is a joint United States Office of Education and National Institute of Education effort. It is the responsibility of the JDRP to identify projects that they believe to be successful and worthy of dissemination.

Diffusion. Diffusion is the process of sharing a validated Title IV-C project with schools outside the originating site.

Background

Innovation

"An innovation is an idea, practice, or object perceived as new by an individual" (Rogers & Shoemaker, 1971, p. 19). The concept of innovation is not new to education. Socrates and Dewey were both educational innovators in their time. In contemporary education, innovation gained a new importance during the Johnson administration. Federal involvement in the process of educational innovation was established under Title III of the Elementary and Secondary Education Act.

> ESEA Title III, programs for supplementary centers and services, was part of the landmark Elementary and Secondary Education Act of 1965. This act truly initiated Federal involvement in the support and direction of public education. Although much Constitutional debate took place concerning the proper role of federal and state governments, the leadership of the Johnson administration resulted in major revision of the Federal role in education.

> Title III of the act provided funds, on a competitive basis to local schools, for the development of innovative programs. These programs which were usually three-year grants, addressed a wide variety of topic areas.
>
> The education amendments of 1974 brought about the consolidation of numerous categorical programs. Title IV, Part C, was created through the consolidation of ESEA Title III with programs addressing dropout prevention and health and nutrition. While the word "innovative" has become unpopular, Title IV, Part C was designed to provide funds, through competitive grants, for the development of programs to improve educational practices. An increased emphasis has been placed on the evaluation of the effectiveness of these grants, with the possibility of a fourth and fifth year of Title IV funds for the dissemination of successful programs. (MDE, 1979).

The passage of the Elementary and Secondary Education Act, made federal funds available to local educational agencies for the development of innovative education projects. The funds are distributed in the form of "seed money." The intent of the legislation is to provide incentive for program development and educational improvement to local school districts who would eventually absorb the costs of the innovation into their regular budgets. Federal involvement with educational innovation became synonomous with Title III and later Title IV-C. ESEA Title IV-C currently supports creative ideas for improving education. Title IV-C projects funded in Massachusetts

deal with new ideas, practices and objects, in accordance with Rogers and Shoemaker's definition.

Ten years ago, USOE officials asked researchers at the Rand Corporation to conduct a study of innovations. While the research continues, an important interim outcome of their work was that an innovative educational project seems to pass through three developmental phases during its evolution. The three phases identified by Rand are:

- Initiation, when LEA officials plan projects and decide which ones to support.
- Implementation, when the project confronts the reality of the institutional setting and project plans must be translated into practice. We hypothesize that effective implementation requires mutual adaption between the project as planned and the institutional setting, in which each must adjust to the demands of the other.
- Outcomes or Incorporation, when the innovative practive loses its "special project" status and becomes part of the routinized behavior of the district. In this phase the project may be continued in whole or in part as a result of deliberate district decisions, or aspects of the innovation may be incorporated by individual teachers with or without formal district support. (1975, Vol. IV, p. 3)

Within this system, as an educational innovation passes through developmental phases it becomes important to manage the innovation accordingly throughout and beyond the federal funding period planning for implementation, evaluation, validation, and diffusion.

The important ESEA legislation assists local school districts in finding new answers to existing problems. Local commitment to change here in Massachusetts and the promotion of quality education is synonymous with ESEA Title IV-C.

Evaluation

The pressure on eduators to evaluate the impact of their educational programs has increased during the last decade. Education has come under scrutiny and budgets have been cut as a result of inflation and taxpayer resistence. Parents and taxpayers demand accountability from school systems, they call for documentation of successful outcomes and expect the results of their programs to be publicized. In addition, state and federal governments are beginning to ask local school districts to state their objectives and measure their effectiveness. Such public pressure may cause an educational evaluation to be seen as producing more problems than gains for the local school system.

When the United States Congress passed the Elementary and Secondary Education Act, for the first time in history, legislation required that educational projects be evaluated. Suddenly, educators were required to evaluate their innovations. On the whole, educators had little expertise in educational evaluation and research

and were generally unprepared to implement the law. The resulting evaluations were inadequate and of little use to anyone. Over the next decade advances were quickly made in the field of educational evaluation. Despite the methodological advances few practitioners on the local level were able to effectively evaluate their programs. In the case of Title III and Title IV-C in Massachusetts this resulted in innovative education programs that appeared on the surface to be providing a great service to America's youth but lacked evidence of student effectiveness and thus could not prove the worth of the innovation. Generally, an innovative project could not be validated because the evaluation design did not provide sufficient information about the student outcomes of the innovation.

The intention of Title IV-C funds is to diffuse programs deemed successful through validation to other communities for a fraction of the initial federal investment. When success is equated with validation, an enormous importance must be placed on program evaluation, which is the vehicle that provides the evidence of program effectiveness for validation. In this respect, program evaluation becomes an integral aspect of project management, and in fact, one aspect that must be carefully attended to if program success is to be determined at the conclusion of the federal funding period.

Program evaluation must be systematically injected into the everyday management of the project. Evaluation must begin as soon as the federal funding process begins with the operationalization of goals and objectives, and continue to grow with the project to provide the information necessary to plan, implement, and validate the project. In order to accomplish this, directors must be good consumers of evaluation. In other words, these directors must know what constitutes a good program evaluation for validation in order to manage it effectively. The program evaluator must be skilled and experienced in the various aspects of educational evaluation. The evaluator must be able to control for the many specific evaluative problems that plague educational innovations. Small sample size, lack of appropriate standardized instrumentation, and the inability to randomize children are a few examples.

The fact that not one Title IV-C project in the Commonwelath of Massachusetts has been validated by the Joint Dissemination Review Panel indicates that program evaluations have not been sufficiently sensitive to provide information in support of program success or failure resulting in a default of non-validation. Why Title IV-C program evaluations in Massachusetts do not satisfy federal validation requirements as well as solutions to the problem are presented herein.

Validation

The United States government has invested billions of dollars in the development of innovations in order to improve the quality of education in America. Most subjects in the elementary and secondary curriculum have been enriched by innovations developed with federal funds. The projects attempt to meet educational needs that have gone unrecognized. Because the federal government pays for the development of these innovations through the United States Office of Education (USOE) and the National Institute of Education (NIE) (the two agencies that comprise the Education division of the Department of Health, Education and Welfare), this does not insure that the innovations will always be effective, significant, sound, and cost-effective. Some projects are successful in that they attain their objectives, others are not. NIE and USOE are left with the quality control problem of separating the successful programs from the unsuccessful. A quality control mechanism was put in place in the form of a review process. From the beginning, the quality control review process was conceptualized as a joint USOE-NIE activity. However, in 1973, high-level USOE program officers formed the Dissemination Review Panel. NIE implemented a similar review process. The function of both panels was to review claims of effectiveness presented by program officers from USOE or NIE.

In 1975 the two panels were combined to form a joint USOE-NIE activity. The resulting panel is the Joint Dissemination Review Panel (JDRP).

The Joint Dissemination Review Panel is comprised of eleven staff members from USOE and eleven members from NIE. Members of the JDRP are nominated by the Commissioner of Education and the Director of NIE from among their staffs. The panel members are chosen on the basis of their qualifications in evaluation and social science methodology, the programs they represent, and their position within the agency.

The JDRP serves only the Education Division of the Department of Health, Education and Welfare. It is the responsibility of USOE and NIE program officers to identify projects that they believe to be successful and worthy of dissemination. In the case of Title IV-C, the USOE program officer submits a worthy project to the JDRP for validation. Only those innovations which have been developed with federal funds may be reviewed by the JDRP. The program officer submits the application, but the project director and project evaluator usually prepare the application according to JDRP guidelines. The program officer conducts a pre-review panel of the submission. The pre-review panel will examine the submission for accuracy, consistency, completeness, compliance with regulations, social fairness, length (the submission may

be only 10 pages in length), the absence of harmful elements, and educational significance. Before a project is worthy of the recommendation of the federal program officer, the project should be approved in the originating state by the Identification, Validation, Dissemination Process (IVD) or by a state approved process such as the Massachusetts Validation Process. Once a project is validated in Massachusetts for diffusion, the state Title IV-C coordinator may recommend the project to the federal program officer, who in turn conducts a pre-review panel and may recommend submission to the JDRP.

The criteria for judging effectiveness varies from the state level to the federal level. On the federal level the criteria for validation arepresented in the IDEABOOK, The Joint Diseemination Review Panel (1977). Following are the seven items listed in the above as guidelines for judging the effectiveness of a project:

- Interpretability of measures: Evidence that the quantitative measures are reliable and valid indicators of the effects claimed.
- Credibility of evidence: Who collected and analyzed the data, what assurances are there that the findings are objective?
- Evidence of impact: What is the evidence that something happened? What are the effects claimed for the intervention?
- Evidence of statistical reliability of the effects: What is the evidence that the effects happened often enough and with sufficient reliability to be likely to happen again under similar circumstances?

- Evidence that the effects are educationally meaningful: What is the evidence that the effects are large enough, powerful enough, or important enough to be educationally meaningful, regardless of their statistical significance.

- Evidence that the effects are attributable to the intervention: Can alternative explanations such as practice effects, maturation, selection of superior treatment groups, etc., be ruled out?

- Evidence of generalizability to the populations for which the product or practice is intended: Evidence that the product or practice has been tested widely enough and under sufficient diverse circumstances to give assurance that the effects claimed may be similar when the product or practice is used elsewhere for the populations intended. (p. 75)

The identification, validation, dissemination process (IVD) was developed by several national groups for the purpose of validating various federally funded programs. The process is used by some states in place of a state validation process such as the Massachusetts Validation Process. The review process has a research orientation. Projects which successfully attain validation through the IVD are usually encouraged to attempt validation by the Joint Dissemination Review Panel.

The Massachusetts Validation Process was developed by the ESEA Title IV-C Coordinator and staff and is used to validate worthy innovations in Massachusetts. The Process as initiated in 1978-1979 focused upon three criteria. The criterion are presented in Massachusetts

Validation Process, Overview (1978). The process focuses upon the following criteria:

1. Evidence of Effectiveness: Supporting evidence is provided to show that the attainment of the major objective(s) can be attributed to the project activities.

2. Exportability: Information is provided to demonstrate that it is reasonable to transport the program or practice to other school districts and that it can be adopted or adapted by other school districts.

3. Economic Efficiency: Sufficient information is provided describing needed costs of start-up, operation and management, and the population to be served, which, when combined with evidence of effectiveness and exportability, will assist an interested school district to make an informed decision about adoption or adaptation of the program of practice. (p. 7)

The Massachusetts validation process is accomplished by an on-site visit of one day by a validation team.

Following an on-site visit the innovation will be placed in one of three categories:

1. Validated as an exemplary model for state diffusion.

2. Validated for state information dissemination.

3. Not validated.

When a project is validated for diffusion in Massachusetts, the Massachusetts Department of Education may support an attempt for validation through the JDRP process.

Diffusion

In Massachusetts when a Title IV-C project is validated for diffusion it achieves exemplary project status. The project may diffuse to other sites in Massachusetts. A Title IV-C project receives three years of development money. Once a project is validated for diffusion by the Massachusetts Validation Process it receives an additional 10% funding in its third year to package products in preparation for diffusion. In addition, the project will receive a fourth year of funding for diffusion activities. Local education agencies wishing to adopt a Title IV-C innovation that is validated for diffusion may apply for an adopter grant in any state funding cycle.

On a federal level the task is more difficult. At this level a project is not moving from town to town, it is moving from state to state. NIE and USOE have initiated unique diffusion activities. NIE funds several nationwide projects designed to develop an effective strategy for dissemination. USOE has established the National Diffusion Network (NDN) to serve USOE funded innovations that have passed JDRP validation. The National Diffusion Network is a system that presents information and technical assistance to local school systems so that they may adopt JDRP validated projects.

A project must be validated by the Joint Dissemination Review Panel in order to be disseminated by USOE and the National Diffusion Network. The agency will not endorse or advocate the use of an innovation that has not been validated.

A project that has been validated may attain developer-demonstrator status from the National Diffusion Network if it presents a proposal for dissemination that merits federal support. Funds are limited, and USOE does not grant all JDRP validated projects development-demonstrator status. Projects are chosen that: (1) meet a pressing national education need and (2) have sound dissemination plans.

Assumptions of the Study

For the purposes of the study, it was assumed by the investigator that:

1. Project directors and validators responded honestly and candidly to questions on both the Massachusetts Statewide Assessment Questionnaires and the Validator Questionnaires.

2. Evaluation experts as well as the participants of the Massachusetts Evaluation Needs Assessment Conference, advised and critiqued with professional and ethical integrity.

3. The interviewer (other than the study investigator), who conducted interviews with Massachusetts Title IV-C personnel, was sensitive to the subjects' responses, and did not misrepresent the subject.

4. The Massachusetts Title IV-C personnel who were interviewed, responded candidly and honestly to all questions.

Limitations of the Study

This dissertation presents an evaluation of the evaluation efforts of Title IV-C projects in Massachusetts. In this sense, the study was a meta-evaluation of project evaluation efforts as they relate to validation. The study is limited in that it was not a research study involving experimental and control groups. Rather, it was a descriptive study.

As described in Chapter III every effort was made to insure the acquisition of valid data, the systematic analyses of these data, and the objective presentation of the data. Nevertheless, the generalizability of outcomes of this study to other state settings must be tempered due to the uncertain quality of the data upon which this study rests.

Significance of the Study

Validation, as administered by the Joint Dissemination Review Panel, is a major intention of federal education agencies. It is a focal point of the new

Title IV legislation which increases emphasis on improvement of education practices through validation and diffusion. Validation has also become a priority of the Massachusetts Department of Education as demonstrated by the 1978-1979 operational plan of the Commissioner of Education. To date, the success of Massachusetts projects at validation has been poor. This lack of success with validation is not unique to Massachusetts. The problem is wide-spread as is the cause, poor project evaluation. Project evaluations must be strengthened so that conclusive evidence of student change can be gathered. A study of the evaluation of Title IV-C innovative education projects in Massachusetts should prove important to researchers, evaluators, validators and administrators in Massachusetts, in other states, and on the federal level, and to Title IV-C personnel as well as to those involved with other types of discretionary grants of an innovative nature. This study has already received much attention from USOE personnel.

The study is unique in the sense that the evaluation of innovative education will be examined in developmental terms. In doing so, project evaluation will mesh with project administration, resulting in a less threatening situation in the local school system.

Studying the evaluation of Title IV-C projects in Massachusetts may have implications for future innovative projects. Results of the study may be of interest to those who would like to improve a validation success rate in other states, as well as those who would be interested in an administrative system which may improve the quality control of the evaluation of innovative education.

CHAPTER II

REVIEW OF THE LITERATURE

Introduction

The purpose of the review of the literature is to establish a theoretical base for the study. To meet this end, the review of the literature is divided into three sections. Section One examines the changes in program evaluation from the earliest recorded evidence of evaluation through the 1970's. Program evaluation is reviewed and the characteristics of a comprehensive educational evaluation are summarized. This review leads to a detailed description of educational evaluation in Sections Two and Three. Each of the characteristics of the comprehensive program evaluation are reviewed and summarized. Section Two analyzes the current state of formative evaluation with an emphasis on design. Finally, Section Three deals with the state of the art of summative evaluation and examines current trends in research design, instrumentation, reliability, validity, and data analysis techniques. Each of the evaluation characteristics presented in this chapter are of the utmost importance to program evaluation for validation and must be meticulously attended to by the successful evaluator.

In total, this review of the literature emphasizes the changes in educational evaluation over the past several decades. Special attention is given to the evaluation of innovative education. Chapter V deals with the application of the evaluative theory presented in this chapter.

Evaluation

Program Evaluation Milestones

Evaluation has been practiced for thousands of years. DuBois (1970) suggests that evaluation was evident in China as early as 2000 B.C., and that Socrates verbally evaluated his students. In the United States, early evidence of program evaluation was a comparative study of the spelling performance of 33,000 students in schools located in the Northeast quarter of the United States. Results were reported by Joseph Mayer Rice in 1898 (Worthen & Sanders, 1973).

The first utilization of standardized instrumentation in program evaluation was by William Learned and Ben Wood in the 1929-1938 "Pennsylvania Study" (Learned & Hawkes, 1940). Learned and Wood evaluated the academic performance of high school and college graduates. Their evaluation included a "scheme for measuring and recording high school progress in terms of

comparable units (as) an effort to try out the effectiveness of repeated objective tests in disclosing that equipment of mastered knowledge which a college would find essential" (p. 25).

Educational evaluation was often conducted in survey form (Thorndike, 1971) in the 1920's and 1930's. Substantial changes occurred by the late 1930's when Smith and Tyler (1942) conducted the Eight Year Study. This evaluation included the revolutionary use of a variety of measurement techniques. The evaluation design used in the Eight Year Study profoundly influenced the design of future evaluations.

In the 1940's formal accrediting agencies for schools and colleges were established (Worthen & Sanders, 1973). The desire on the part of educators to receive accreditation encouraged the use of program evaluation.

Thorndike published the first edition of Educational Measurement in 1951. Early measurement technology flourished during the early half of this century. In the beginning, measurement and evaluation grew as separate entities, until finally, evaluation specialists interwove the technology of each field.

> Historically, formal evaluation has been very closely associated with the measurement tradition in psychology and education. In fact, even today one finds that many writers see little discrimination between the process of measurement and evaluation. This heritage is also evident in the abundance of psychological measurement tools which

> which are used by the evaluation specialist. (Worthen & Sanders, 1973, p. 2)

During the period of 1956-1972 a Taxonomy of Educational Objectives was published in three domains. In 1956, a Taxonomy of Educational Objectives in the cognitive domain was published (Bloom, 1956). The taxonomy includes "objectives which deal with the recall or recognition of knowledge and the development of intellectual abilities and skills" (p. 7). For the first time a "classification of measurable educational outcomes: (p. 10) was available to evaluators. Nearly a decade later a taxonomy of educational objectives in the affective domain was published (Krathwohl, 1964). The taxonomy helped evaluators "become aware of the techniques available for appraising growth of students toward various categories of objectives and for assessing other affective changes" (p. 23). The following decade, a taxonomy of educational objectives in the psychomotor domain (Harrow, 1972) was published. "The specific intent of this text is to provide a functional written taxonomy for the psychomotor domain to be utilized for the classification of observable movement behaviors" (p. 7). These documents provided strategies for educators to identify and classify behaviors.

The Sputnik years, the 1950's and 1960's echoed with cries for curriculum reform. The first comprehensive attempt to evaluate curriculum began in 1956 when the

Physical Science Study Committee (PSSC) was formed (Martin & Pinck, 1966). The PSSC was a group of university and secondary school physics teachers working to develop an improved introductory physics course. The project and subsequent evaluation of it was led by Jerold R. Zacharias of the Massachusetts Institute of Technology. This early attempt at curriculum evaluation resulted, by the Fall of 1960, in the availability of textbooks, laboratory guidebooks, apparatus, tests, films, and teacher's guidebooks. The evaluation included the committee's internal assessment of its objectives as well as examination of students performance on PSSC achievement tests. Comparisons of student performance in the PSSC course and other physics courses was not undertaken. Even though the PSSC effort was the first major attempt at curriculum reform through evaluation, this effort became subject to criticism in later years.

In 1963, Cronbach (1973) expressed the opinion that the greatest service an evaluation can perform is to identify aspects of a course where revision is desirable.

> I am becoming convinced that some techniques and habits of thought of the evaluation specialist are ill-suited to current curriculum studies. Old habits of thought and long-established techniques are poor guides to the evaluation required for course improvement.
>
> Course evaluation should ascertain what changes a course produces and should identify aspects of the course

> that need revision. The outcomes observed should include general outcomes ranging far beyond the content of the curriculum itself; attitudes, career choices, general understandings and intellectual powers, and aptitude for further learning in the field. Analysis of performance on single items or types of problems is no more informative than analysis of composite scores. It is not necessary or desirable to give the same test to all pupils, rather, as many questions as possible should be given, each to a different moderate-sized sample of pupils. Costly techniques such as interviews and essay tests can profitably be applied to samples of pupils, whereas testing everyone would be out of the question.
>
> Asking the right question about educational outcomes can do much to improve educational effectiveness. Even if the right data are collected, evaluation will have contributed too little if it only places a seal of approval on certain couses and casts others into disfavor. Evaluation is a fundamental part of curriculum development, not an appendage. Its job is to collect facts the course developer can and will use to do a better job, and facts from which a deeper understanding of the educational process will emerge. (p. 58)

In the early 1970's a program evaluation that focused on curriculum alone was considered to be inadequate. Evaluation design was expanded to a decision-management approach.

The CIPP (Context, Input, Process, Product) model approached evaluation for decision-making. Stuffelbeam defined evaluation as "the process of delineating, obtaining, and providing useful information for judging

decision alternatives" (1971, p. 120). Stuffelbeam opened new doors to evaluation when he suggested that internal validity, external validity, reliability, objectivity, relevance, importance, scope, credibility, timeliness and effeciency were important components of evaluation.

Hutchenson and Benedict (1973) identified many gaps in evaluation theory and documented a need for methodological research in educational evaluation.

> Beginning with a more comprehensive and more utilitarian definition of the purpose of evaluation, namely to provide data for decision making, (the authors) have developed perceptive, not merely descriptive, procedures for educational evaluation. In fact, the only legitimate function of this evaluation methodology is to provide data to decision makers for their decision making purposes. (p. 4)

The evaluation methodology includes further implications overlooked by other "decision-oriented" models. Three "user" criteria emerged for evaluation practice: (1) efficiency: An evaluation is efficient to the extent that it provides only that data which a decision maker actually uses; (2) completeness: An evaluation is complete to the extent that it provides all the data needed by a decision maker; and (3) focus: An evaluation is focused to the extent it provides all the data for the decision makers highest priority needs.

These three "user" criteria have counterparts on the evaluation methodology level, i.e., "evaluator criteria": (1) efficiency for the evaluator implies a continuing high degree of contact with the decision maker and continuing review by the decision maker; (2) completeness implies that the methodology is tested for completeness as to the decision makers' needs on a continuing basis; and (3) focus implies that methodology use decision maker priorities at every stage, rather than the evaluator's or someone else's. This unorthodox method of program evaluation for decision maing is unique in that it includes procedures for evaluators to measure the effectiveness of their own work.

The unfolding of historical concern over educational evaluation led to early attempts at evaluation. Worthen and Sanders noted that:

> It was in this context that the United States Congress began its deliberations on the proposed Elementary and Secondary Education Act of 1965. Robert F. Kennedy was among the senators who forcefully insisted that the ESEA carry a proviso requiring educators to be accountable for the federal monies they receive and to file an evaluation report for each grant showing what effects had resulted from the expenditure of the federal funds . . . these efforts led to the first major evaluation mandate issued . . . this meant that thousands of educators were for the first time required to spend their time evaluating their own efforts . . . educators

> were unprepared to implement the new mandate effectively . . . many of the evaluators were inadequate . . . the results of these evaluations were not of much use to anyone. (Worthen & Sanders, 1973, p. 6)

Title III and later Title IV part C of the Elementary and Secondary Education Act made federal funds available to local educational agencies for the development of innovative education projects. The funds are distributed in the form of "seed money." The intent of the legislation is to provide incentive for program development and educational improvement to local school districts who would eventually absorb the costs of the innovation into their regular budgets. For the first time in history, federal legislation required that educational projects be evaluated. Suddenly, educators were required to evaluate their federally-funded innovations. On the whole, educators had little expertise in educational evaluation and research and were generally unprepared to implement the law. The resulting evaluations were inadequate and of little use to anyone. The plea for responsible educational evaluation was echoed, and a heavy emphasis continued to be placed on curriculum evaluation.

Scriven pointed out that "the main deficiency (in educational evaluation) is trained evaluation manpower" (p. 103). In addition, Scriven suggested that educators were not investing adequate resources in the evaluation process.

> The educational profession is suffering from a completely inappropriate conception of the cost scale for educational research. To develop a new automobile engine or a rocket engine is a very, very expensive business despite the extreme constancy in the properties of physical substances. When we are dealing with a teaching instrument such as a new curriculum or classroom procedure, with its extreme dependence upon highly variable operators and recipients, we must expect considerably more expense. The social pay-off is enormously more important, and this society can, in the long run, afford the expense. (p. 103)

In 1967 Robert E. Stake argued that just as educators differ, evaluation methods differ which allows each educator to keep his own perspective of education rather than examine the full countenance of evaluation.

> Formal evaluation of education is recognized by its dependence on check-lists, structured visitation by peers, controlled comparisons, and standardized testing of students. Some of these techniques have long histories of successful use. Unfortunately, when planning an evaluation, few educators consider even these four. The more common notion is to evaluate informally to ask the opinion of the instructor, to ponder the logic of the program, or to consider the reputation of the advocates. Seldom do we find a search for relevant research reports or for behavioral data pertinent to the ultimate curricular decisions. School officials cannot yet reuse curriculum on rational grounds, and the needed evaluation is not underway . . . In our data banks we should document the causes and effects, the congruence of intent and accomplishment, and the panorama of judgments of those concerned. Such records should be kept to promote educational action, not obstruct it. The countenance of evaluation should be one that leads to decision-making, not to trouble-making. (Stake, 1967, p. 539)

Stake suggested that an evaluation should include descriptive and judgmental data and provide immediate relative answers.

Cronbach (1963), suggested a revolutionary two-stage approach to program evaluation.

> The approaches to evaluation include process studies, proficiency measures, attitude measures, and follow-up studies. A process study is concerned with events taking place in the classroom, proficiency and attitude measures with changes observed in pupils, and follow-up studies (are concerned) with the later careers of those who participated in the course.
>
> The follow-up study comes closest to observing ultimate educational contributions, but the completion of such a study is so far removed in time from the initial instruction that it is of minor value in improving the course of explaining its effects. The follow-up study differs strikingly from the other types of evaluation study in one respect. I have already expressed the view that evaluation should be primarily concerned with the effects of the course under study rather than with comparisons of courses. That is to say, I would emphasize departures of attained effectiveness of different parts of the course, and differences from item to item; all these suggest places where the course could be strengthened. But this view cannot be applied to the follow-up study, which appraises effects of the course as a whole and which has very little meaning unless outcomes can be compared with some sort of base rate. Suppose we find that 65 percent of the boys graduating from an experimental curriculum enroll as scientific and technical majors in college. We cannot judge whether this is a high or low figure save by comparing it with the rate among boys who have not had the course. In a follow-up study, it is necessary to

> obtain data on a control group equated at least crudely to the experimental cases on the obvious demographic variables. (p. 678)

Scriven (1967) took the Cronbach approach to program evaluation one step further with the argument that program evaluation should be formative and summative in nature. Formulative evaluation is associated with the ongoing improvement of curriculum. Summative evaluation provides final evidence of the worth of the finished curriculum, as well as justification of the expense of adoption by another school system.

This important development has had a major impact on the future of program evaluation. Rossi (1979) distinguishes between summative and formative evaluation yet clearly states that the comprehensive (program) evaluation is one that includes both formative evaluation and summative evaluation. In addition, Bloom (1971) has suggested that a program evaluation "should be both formative and summative in its scope" (p. 20). Formative evaluation and summative evaluation will be examined individually later in this chapter.

Alken (1973), suggested that each evaluation should be designed according to the unique aspects of the program in question.

> Evaluation is the process of ascertaining the decision areas of concern, selecting appropriate information, and

> collecting and analyzing information in order to report a summary data. (p. 150)

Recent literature suggests that "program evaluation is continuing and ongoing. It occurs at the start, during and after a program has been run. One may consider evaluation as the nucleus of the program, for it interacts with the program's needs assessment, its statement of goals and objectives, and program planning and implementation" (CSDE, 1977).

The evaluation theorists of the 1960's and early 1970's created designs that brought educational evaluation out of the dark ages, "Yet, despite these trends toward accountability, only one tiny fraction of the educational programs operating at any level have been evaluated in any but the most cursory fashion, if indeed at all" (Worthen & Sanders, 1973, p. 1).

The methodology of evaluation matured from the survey format of the 1930's to the sophistication of the 1970's. Yet, despite the newly developed evaluation methodologies, useful information about them was scarce at best. The connection between the theory of evaluation and the practice of evaluation was not being made in the local school system. This resulted in inadequate evaluation. In the case of Title III and Title IV-C this resulted in innovative education programs that appeared

on the surface to be providing a great service to America's youth but lacked evidence of student effectiveness and thus could not prove the worth of the innovation. Generally, an innovative project could not be validated because the evaluation design did not provide significant information about the students outcomes of the innovation.

Formative Evaluation

Educational evaluation can be broken down into two specific categories, formative evaluation and summative evaluation.

> Formative evaluation takes place during the development of a program or instructional unit. It is concerned with fine tuning the implementation process and measuring learner progress toward the attainment of specified objectives. Thus, formative evaluation provides the decision-maker with information during the course of program development and execution for possible mid-course corrections to help assure that the program objectives are eventually met in an effective and economical fashion. (CSDE, 1977, p. A6)

The preliminary work in developing a formative evaluation plan requires consideration of all of the steps involved in the evaluation process. "Evaluation design is essentially a systematic approach to the task of gathering information to answer questions or make decisions" (CSDE, 1977, p. B1). The formative evaluation

design should include project goals, activities, evidence of program merit, evaluation questions, information collection techniques, schedule, and personnel (Steiger, 1976).

An educational project contains a series of undifferentiated ends-goals and means-goals (Walker & Wolf, 1979). Formative evaluation is for decision-making purposes, program management, and program improvement and is concerned with means-goals. "Means-goals are important for measuring project outcomes. Means-goals will fit this sentence: During the program, students, or parents, or teachers, will take part in. . . ." (p. 10).

An activity is a means of achieving a goal (CSDE, 1977, p. B3). The activities that will be conducted to realize each goal, should be listed and included in the design (Steiger, 1976). "The actual evidence that a program has [formative] merit will take the form of statements, events, objects, and observations that testify to its quality. At least one indication of program merit must be identified for each program and each activity" (CSDE, 1977, p. B5).

Information collection techniques may consist of norm-referenced tests, criterion-referenced tests, or home grown tests. Home grown tests are frequently used in formative evaluation, because one is simply gathering

data for program improvement rather than assessing change in the target population (summative evaluation) which would require a valid and reliable test. In addition, questionnaires, interviews, observations, rating sheets, logs, record summary forms, and narrative reports are frequently used in formative evaluation (CSDE, 1977). One should consider the following criteria when choosing formative information collection techniques: (1) Does the instrument adequately measure what it should?; (2) Is the instrument appropriate for the particular population?; (3) Is the instrument easy to administer and score?; and (4) Is the cost of the instrument, its administration and its scoring, reasonable and within the budget? (p. B3).

Formative evaluation requires collecting and sharing information for program improvement. "While a program is being installed, the formative evaluation provides the program planners and staff with information to help adjust it to the setting and improve it" (Morris & Fitz-Gibbon, 1978, p. 9). Formative evaluation is an internal evaluation and should be conducted by the project staff.

Formative evaluation should be conducted when an innovation is in the early phases of its development. The formative evaluation provides the staff of the Title IV-C program with information that would aid in seeking support and commitment for the project. In addition,

formative data would allow the project director to make decisions to modify the project so that adoption into the school setting may take place. Formative evaluation is management oriented and should be conducted by project staff thus freeing up evaluation funds for the more rigorous summative evaluation.

Summative Evaluation

Summative evaluation is more rigorous than formative evaluation. Rossi (1979) has defined summative evaluation or evaluative research as an "assessment of the extent to which a program causes change in the target population" (p. 16). Walker and Wolf (1979) have differentiated between the two in the following list (p. vi).

Summative	Formative
At the end of the project, more formal	Throughout the project more informal
Program documentation	Program monitoring
Produces a summary statement about effectiveness	Produces statements that will encourage dialogue about the program and lead to program growth
Uses multiple instrumentation with high degree of reliability and validity	Uses instruments that provide information about program, attitude, or achievement for program development; develops or tests instruments for use in summative evaluation

"Summative evaluation differs from formative evaluation in its timing and audience. When a program has passed its developmental stage and is functioning as intended, it is ready to be summarily described and perhaps judged" (p. 9). The summative evaluation requires greater objectivity and therefore should be conducted by an outsider who is trained in measurement theory.

Nunnally (1975), defines summative evaluation as being "generally concerned with the effectiveness of programs of social improvement" (p. 101).

Federally funded projects such as Title IV-C projects may be reviewed for worthiness by the Joint Dissemination Review Panel (Tallmadge, 1977). The Joint Dissemination Review Panel (JDRP) was established by the Education Division of the Department of Health, Education and Welfare in 1972. The JDRP meets periodically to review the evidence of effectiveness submitted for a wide variety of educational products and practices. The JDRP has established strict evaluation standards. In order to meet the criteria established by the JDRP a summative evaluator must consider the appropriateness of the evaluation design in the areas of research design, instrumentation, reliability, validity, sampling, randomization and statistics.

Research Design

The application of an experimental or research design to educational programs was examined by Campbell and Stanley (1963):

> . . . across the centuries many different approaches [to teaching] have been tried, if some approaches have worked better than others, and if those which worked better have therefore, to some extent, been more persistently practiced by their originators, or imitated by others, or taught to apprentices, then the customs which have emerged may represent a valuable and tested subset of all possible practices.
>
> But the selective, cutting edge of this process of evolution is very imprecise in the natural setting. The conditions of observation, both physical and psychological are far from optimal. What survives or is retained is determined to a large extent by pure chance. Experimentation enters on this point as the means of sharpening the relevance of the testing, probing, selection process. Experimentation thus is not in itself viewed as a source of ideas necessarily contradictory to traditional wisdom. (p. 4)

Further, Campbell and Stanley (1963) have suggested that the experiment is "the only way of verifying educational improvements, and is the only way of establishing a cumulative tradition in which improvements can be introduced without the danger of a faddish discard of old wisdom in favor of inferior novelties" (p. 2).

The quality of an experimental or research design can be evaluated in terms of two criteria, internal validity and external validity. Campbell and Stanley (1963), distinguish between the two classes of validity.

> Fundamental to this listing is a distinction between internal validity and external validity. Internal validity is the basic minimum without which any experiment is uninterpretable: Did in fact the experimental treatments make a difference in this specific experimental instance? External validity asks the question of generalizability: To what populations, settings, treatment variables, and measurement variables can this effect be generalized? Both types of criteria are frequently at odds in that features increasing one may jeopardize the other. While internal validity is the *sine qua non*, and while the question of external validity, like the question of inductive inference, is never completely answerable the selection of designs strong in both types of validity is obviously our ideal. (p. 5)

It is externally important that the questions of internal and external validity be attended to in the summative evaluation of Title IV-C projects. Campbell and Stanley suggest that "this is particularly the case for (evaluation) of teaching, in which generalization to applied settings of known character is the disideratum" (p. 5).

Campbell and Stanley suggest that there are twelve factors that jeopardize the validity of various experimental designs.

Relevant to internal validity, Campbell and Stanley have identified eight different classes of extraneous variables; these variables, if not controlled in the experimental design, might produce effects confounded with the effect of the experimental stimulus. They represent the effects of:

1. History, the specific events occurring between the first and second measurement in addition to the experimental variable.

2. Maturation, processes within the respondents operating as a function of the passage of time per se (not specific to the particular events), including growing older, growing hungrier, growing more tired, and the like.

3. Testing, the effects of taking a test upon the scores of a second testing.

4. Instrumentation, in which changes in calibration of a measuring instrument or changes in the observers or scorers used may produce changes in the obtained measurements.

5. Statistical regression, operating where groups have been selected on the basis of their extreme scores.

6. Biases resulting in differential selection of respondents for the comparison groups.

7. Experimental mortality, or differential loss of respondents from the comparison groups.

8. Selection-maturation interaction, etc., which in certain of the multiple-group quasi-experimental designs, is confounded with, i.e., might be mistaken for, the effect of the experimental variable. (p. 5)

The factors jeopardizing external validity or representatives are:

9. The reactive or interaction effect of testing in which a pretest might increase or decrease the respondent's sensitivity or responsiveness to the experimental variable and thus make the results obtained for a pretested population unrepresentative of the effects of the experimental variable for the unpretested universe from which the experimental respondents were selected.

10. The interaction effects of selection biases and the experimental variable.

11. Reactive effects of experimental arrangements, which would preclude generalization about the effects of the experimental variable upon persons being exposed to it in nonexperimental settings.

12. Multiple-treatment interference, likely to occur whenever multiple treatments are applied to the same respondents, because the effects of prior treatments are not usually erasable. (p. 5)

Campbell and Stanley examine the benefits and problems associated with three Pre-experimental Designs, three True Experimental Designs, and fourteen Quasi-experimental Designs. A summary of the positive and negative aspects of the sixteen designs are presented in Talbe 1 (Swaminathan, 1979, p. 63 and p. 64).

Nunnally (1975), has identified a major problem in the measures employed to evaluate a program concerned with educational improvement.

> A problem that is encountered frequently (in summative evaluation) is that pretest measures interact with treatment conditions in a program being evaluated. To prevent such reactivity, a better research design is one in which tests are given only at the end of a training program and in which the group is compared with a control group. The only major exception

Table 1

Sources of Invalidity

Design	Sources of Invalidity										
	Internal								External		
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements
Pre-Experimental Designs:											
1. One-group Pretest-Posttest Design *O X O*	No	No	No	No	?	Yes	Yes	No	No	No	?
2. Static-group Comparison *X O* --- *O*	Yes	?	Yes	Yes	Yes	No	No	No		No	
True Experimental Designs:											
3. Pretest-Posttest Control Group Design *R O X O* *R O O*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	?	?
4. Posttest-Only Control Group Design	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	?	?

Table 1 (continued)

Design	Sources of Invalidity										
	Internal								External		
	History	Maturation	Testing	Instrumentation	Regression	Selection	Mortality	Interaction of Selection and Maturitation, etc.	Interaction of Testing and X	Interaction of Selection and X	Reactive Arrangements
Quasi-Experimental Designs:											
5. Nonequivalent Control Group Design *O X O* --- *O O*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	?	?
6. Separate-Sample Pretest-Posttest Design *R O (X)* *R X O*	No	No	Yes	?	Yes	Yes	No	No	Yes	Yes	Yes
7. Separate-Sample Pretest-Posttest Control Group Design *R O (X) O* *R X O* --- *R O* *R O*	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes
8. Time Series	No	Yes	Yes	?	Yes	Yes	Yes	Yes	No	?	?

Adapted from Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago, IL: Rand McNally, 1966.

NOTE: In this table, "Yes" indicates that the factor is controlled, a "No" indicates that it is a weakness, a "?" indicates a possible source of concern, and a blank indicates that the factor is not relevant.

> to the foregoing principle is in a situation where the number of subjects necessarily is quite small, and thus every ounce of statistical muscle is needed. The major advantage of a repeated measures design in which the same subjects are administered a pretest and a posttest is that the statistical power of the comparisons frequently are substantially higher than in employing a between groups design, in which a treatment group is compared with a control group. (p. 103)

It is extremely important to consider the degree to which pretest measures interact with treatment conditions when planning the research design associated with the summative evaluation of a Title IV-C program. The evaluator must rule out variables associated with internal validity and be certain that the implementation of the innovation alone caused the change in the children served by the project.

Campbell and Erlebacher (1975) suggest that "true" experiments in the field setting are more "quasi" than those in the laboratory, and those in the laboratory more "quasi" than published reports and statistical treatments indicate.

Nunnally (1975) has defined a quasi-experimental design as one "in which some potentially important confounding variables are not controlled" (p. 125). Nunnally suggests that "what one does in this situation is to obtain extra information (over that which would be obtained from a true experimental design) in order to

weave a net of circumstantial evidence regarding the "reality" of observed findings. Evaluation usually takes place in the crucible of ongoing important life activities of individuals, and consequently it frequently is either unethical or unfeasible to randomly divide subjects up for treatments or to do other things that would be needed to obtain precise experimental control" (p. 125).

In addition, Nunnally and Wilson (1975) caution that "mentioning these problems, however, should not imply a derogation of evaluation or even hint that such research should be discontinued. It is far too easy for the pure methodologist to point up the stringent requirements for pristine research and to induce guilt feelings in the poor souls who have to do the best they can in evaluating programs of social action. The social problems involved will not go away simply because they pose problems for research" (p. 231).

An important consideration in evaluation is the establishment of the experimental and control groups. Campbell and Erlebacher (1975), suggest that evaluations with randomized assignment to treatments are generally to be preferred where possible. "We believe that any investigator fully attending to the presumptions he is making in using quasi-experimental designs will prefer the random assignment of children to treatments where

this is possible. Social ameliorative changes which are applied or made available to everyone do not readily permit the creation of control groups" (p. 614).

In addition Campbell and Erlebacher suggest:

> There exist in administration, researchers, legislators, and the general public "ethical" reluctances to random assignment. These center around a feeling that the control group is being deprived of a precious medicine it badly needs. But if it be recognized that the supposed boon is in fact in short supply, then it can be seen that the experiment has not increased the number so deprived, but has instead reassigned some of that deprivation so that the ethical value of knowing may be realized. Is randomization as the mode of such reassignment ethically defensible? It might represent an ethical cost [one nonetheless probably worth paying] if all the children in the nation had been rank ordered on need, and those most needy given the compensatory education up to the budgetary and staff limits of the program. But instead, the contrast is with a very haphazard and partially arbitrary process which contains unjust inversions of order of need far more extensive than a randomization experiment involving a few thousand children would entail. These unjust deprivations are normally not forced to our attention, and so do not trouble our ethical sensitivities as does the deprivation of the control group. But there is no genuine ethical contrast here.
>
> Within randomization, there are some designs and stances that may ease any residual ethical burden. For example, the randomization could be limited to the boundary zone, at the least needy edge of those to be treated, and the most needy edge of the untreated. For this narrow band of children, all

> considered as essentially tied at the cutting point on a coarse grained eligibility score, random assignment to treatment and nontreatment could be justified as a tie-breaking process. We would learn about the effects of the program only for a narrow band of talent. We would wonder about its effectiveness for the most disadvantaged. But this would be better than nothing, and better than quasi-experimental information. (p. 615)

Instrumentation

Before it is possible to investigate changes that occur in people with respect to programs of educational improvement, it is necessary to have measures that document such changes. Nunnally and Wilson (1975) define measurement as consisting of "Rules for assigning numbers to objects to represent quantities of attributes" (p. 232).

Tests vary in the amount of knowledge and research required to develop them. The American Psychological Association (APA) has published Standards for Educational and Psychological Tests (1974). A list of standards regarded as essential follow:

1. A test manual should describe fully the development of the test: the rationale, specifications followed in writing items or selecting observations, and procedures and results of item analysis or other research.

2. The test and its manual should be revised at appropriate intervals. The time for revision has arrived whenever changing conditions of use or new research data make any statements in the manual incorrect or misleading.

3. The test, the manual, the record forms, and other accompanying material should help users make correct interpretations of the test results and should warn against common misuses.

4. The test manual should state explicitly the purposes and applications for which the test is recommended.

5. The test manual should describe clearly the psychological, educational, or other reasoning underlying the test and nature of the characteristic it is intended to measure.

6. The test manual should identify any specific qualifications required to administer the test and to interpret it properly.

7. Evidence of validity and reliability, along with other relevant research data, should be presented in support of any claims being made.

8. Test developers or others offering computer services for test interpretation should provide a manual reporting the rationale and evidence in support of computer-based interpretations of scores.

9. The directions for administration should be presented in the test manual with sufficient clarity and emphasis so that the test user can duplicate, and will be encouraged to duplicate, and the administrative conditions under which the norms and the data on reliability and validity were obtained.

10. Instructions should prepare the examinee for the examination: sample material, practice use of answer sheets or punch cards, sample questions, etc., should be provided.

11. The procedures for scoring the test should be presented in the test manual with a maximum of detail and clarity to reduce the likelihood of scoring error.

12. Norms should be published in the test manual at the time of release of the test for operational use.

13. Norms presented in the test manual should refer to defined and clearly described populations. These populations should be the groups with whom users of the test will ordinarily wish to compare the persons tested.

14. In reporting norms, test manuals should use percentiles for one or more appropriate reference groups or standard scores for which the basis is clearly set forth; any exceptional type of score unit should be explained and justified. Measures of central tendency and variability always should be reported.

15. Local norms are more important for many uses of tests than are published norms. A test manual should suggest using local norms in such situations.

16. Derived scales used for reporting scores should be carefully described in the test manual to increase the likelihood of accurate interpretation of scores by both the test interpreter and the examinee.

17. If scales are revised, new forms added, or other changes made, the revised test manual should provide tables of equivalence between the new and the old forms. This provision is particularly important in cases where data are recorded on cumulative records.

18. Where it is expected that a test will be used to assess group rather than individuals (i.e., for schools or programs), normative data based on group summary statistics should be provided.

Testing in program evaluation may be criterion-referenced or norm-referenced.

> Typically, criterion-referenced testing uses as its interpretive frame of reference a specified content domain rather than a specified population of persons. In this respect, it has been contrasted with the

> usual norm-referenced testing, in which an individaul's score is interpreted by comparing it with the scores obtained by others on the same test.
>
> Thus far, criterion-referenced testing has found its major applications in several recent innovations in education. (Anastasi, 1976, p. 97)

Norm-referenced tests should be utilized in the summative evaluation of a project whenever possible. Purchasing and administering these standardized tests is cost effective and permits a comparison in reference to others. However, caution must be exercised in assuring that the test chosen for the program evaluation accurately reflects the goals and objectives of the project. When a standardized norm-referenced test does not match the objectives of the program in question, it may be more appropriate to develop a criterion-referenced test. This test would incluse items that specifically match the objectives of the program and therefore would succeed in providing an accurate measurement of the program that a cost-effective yet insensitive norm-referenced test could not. The development of criterion-referenced tests is costly and time consuming and should most definitely be undertaken when necessary in order to avoid the problems associated with the use of a test which does not truly measure a program, despite budget limitations.

Organizations, able to provide assistance in locating, or developing norm-referenced or criterion-referenced tests exists today. The Educational Testing Service and National Evaluation Systems are two such organizations.

The Educational Testing Service (ETS), founded in 1947 is a nonprofit organization devoted to measurement and research. ETS activities span a broad spectrum reflecting society's heightened need for knowledge that will enhance individual development and institutional excellence and diversity (ETS, p.1).

ETS is engaged in activities ranging from operating the nationwide College Scholarship Service for the College Board to conducting research in such areas as child development and cognitive styles; from evaluating the effectiveness of Sesame Street to developing a computer-based guidance program for college students, called the System of Interactive Guidance and Information.

ETS is best known for its norm-referenced national testing programs. Each year, tests developed and administered by ETS are taken by millions of people at more than 5,000 test centers across the nation and around the world. Some of these tests are designed to predict academic performance. Other tests are used for student guidance and placement, for awarding degree credit for independent or advancement learning, and for occupational licensing, professional certification and professional self-assessment.

National Evaluation Systems (NES) specializes in research and evaluation services for a wide range of clients. The company has offices in Amherst, Massachusetts and Palo Alto, California. NES has pioneered criterion-referenced educational testing and customized evaluation

services which conform precisely to their clients' management and organizational needs.

Standardized instrumentation should be utilized in the evaluation of a Title IV-C project whenever possible. However, if standardized instrumentation is not available and instrumentation must be developed, it is necessary that the test developers write the test and manual in accordance with the APA standards.

A measure of success for a Title IV-C project is validation by the Joint Dissemination Review Panle (JDRP). One criterion utilized by the JDRP in judging the effectiveness of a Title IV-C program is termed "Interpretability of Measures" (Tallmadge, 1977). In other words, the JDRP demands evidence that the quantitative measure are reliable and valid indicators of the effects claimed.

As mandated by the American Psychological Association and the JDRP, evidence for reliability and validity must be provided for each test utilized in the evaluation of a Title IV-C project. A detailed discussion of both reliability and validity follows.

Reliability

Reliability refers to the degree to which the results of testing are attributable to systematic sources of variance (APA, 1974). In other words, a test is reliable to the extent that one can assert confidently

that similar results would be obtained if the test administration was repeated (Sellitz, Wrightsman & Cook, 1976).

Lord and Novick (1968) define the reliability of a test as the squared correlation p^2_{xt} between observed score and true score. That is if observed and true scores could be obtained for every examinee for a test, the squared correlation between observed and true scores is called the reliability coefficient for that test (Allen & Yen, 1979).

There are many ways of defining and interpreting test reliability. The American Psychological Association (1974) identified the following estimates of reliability.

> One method of obtaining the two sets of measurements is by retesting with the identical test. Aside from practical limitations, theoretically, retesting is not ordinarily a desirable method of estimating reliability because the examinee may remember his or her responses to items from one testing to the next. Hence, memory becomes a systematic source of variance and the correlation of the two sets of scores may be higher than the correlation of two sets of scores based on two different but parallel sets of items drawn from the population of items in the same way.
>
> If we want to eliminate memory as a systematic source of variance and to include the effects of item sampling and response variation over time as sources of variance, we may use two sets of items developed or selected

according to the same specifications. These are called parallel forms of the test.

If the effect of content sampling alone is sought without the effects of memory or response variability over time, or if it is not practical to administer two parallel forms with separate time limits, reliability can be estimated from a single administration of an unspeeded test. The test may be divided into two sets of items of equal, or approximately equal, length that are judged by competent authorities to sample as nearly as possible the same functions. Any items based on the same source of data (such as a reading passage) must be assigned to the same set. Then the correlation between scores on the two parallel halves is a matched-half coefficient from which an estimate of the parallel-forms reliability coefficient for the total test may be obtained by a procedure that does not assume that the numbers of items or the variances of the two sets are exactly equal.

Estimates of reliability from a single administration may also be obtained by analysis-of-variance procedures. Such estimates will be spuriously high if the test is speeded or if the items are not independent of each other. On the other hand, for unspeeded tests, such estimates will tend to be lower than matched-half coefficients because they constitute, given certain assumptions, the mean of coefficients obtained by correlating scores on all possible pairs of half of the test.

From the preceding discussion, it is clear that different methods of estimating reliability take account of different sources of error. Thus, from one testing to the other, the result is affected not only by random response of variability and changes in subjects over time but also by differences in administration (especially if different persons administer the test on the two occasions).

> Reliability coefficients based on a single administration of a test exclude response variability over time; these effects on scores do not appear as errors of measurement. Hence, "reliability coefficient" is a generic term. It can be based on various types of evidence; each type of evidence suggests a different meaning. It is essential that any method used to estimate reliability be clearly described.
>
> The estimation of clearly labeled components of score variance is the most informative outcome of a reliability study, both for the test developer wishing to improve the reliability of his instrument and for the user desiring to interpret test scores with maximum understanding. The analysis of score variance calls for the use of an appropriate experimental design. There are many different multivariate designs that can be used in reliability studies; the choice of design for studying a particular test is determined by its intended interpretation and by practical limitations.
>
> It is recommended that test authors describe the meanings of any coefficients they report as accurately and precisely as possible. It is informative to say, for example, "This coefficient indicates the stability of measurement of equivalent scores based on parallel forms of the test administered 7 days apart, without intervening practice or instruction." Although lengthy, such a description is reasonably free from ambiguity. (p. 48)

The person responsible for choosing or developing the instrumentation to be utilized in the summative evaluation of a Title IV-C program must pay special attention to the issue of reliability. Both the Massachusetts Validation Process as well as the Federal Validation

provess (JDRP) require proof that any measurement techniques are consistent in measuring what they are designed to measure. The person responsible should be sure that the reliability score has been appropriately and sufficiently determined.

Validity

Validity refers to the appropriateness of inferences from test scores to other forms of assessment (APA, 1974). The most meaningful measures of the value of a test are its validities (Lord & Novick, 1968). Validity can be assessed in several ways, depending on the test and its intended use. The three major types of validity are content validity, criterion-related validity and construct validity. Determinations of criterion-related validity and construct validity involve the calculation and examination of correlation or other statistics. Content validity, however, does not involve any statistitical calculations.

The American Psychological Association (1974) refers to criterion-related validities in the following way;

> Criterion-related validities apply when one wishes to infer from a test score an individual's most probably standing on some other variable called a criterion. Statements of predictive validity indicate the extent to which an

> individual's future level on the criterion can be predicted from a knowledge of prior test performance; statements of concurrent validity indicate the extent to which the test may be used to estimate and individual's present standing on the criterion. The distinction is important. Predictive validity involves a time interval during which something may happen (e.g., people are trained or gain experience, or are subjected to some treatment). Concurrent validity reflects only the status quo at a particular time. Under appropriate circumstances, data obtained in a concurrent study may be used to estimate the predictive validity of a test. However, concurrent validity should not be used as a substitute for predictive validity without an appropriate supporting rationale. (p. 26)

Content validity is required when the test user wishes to estimate how an individual performs in the universe of situations the test is intended to represent.

> Content validity is most commonly evaluated for tests of skill or knowledge; it may also be appropriate to inquire into the content validities of personality inventoires, behavior checklists, or measures of various aptitudes. The present discussion will be directed toward the more typical case of achievement testing.
>
> To demonstrate the content validity of a set of test scores, one must show that the behaviors demonstrated in testing constitute a representative sample of behaviors to be exhibited in a desired performance domain. Definitions of the performance domain, the users' objectives, and the method of sampling are critical to claims of content validity. An investigation of content validity requires that the test developer or test user specify his objectives and carefully define the performance domain in light of

> those objectives. The definition should ordinarily specify the results of learning rather than the processes by which learning is either sufficiently detailed and organized to show the degree to which component tasks make up the total domain.
>
> It should be clear that content validity is quite different from face validity. Content validity is determined by a set of operations, and one evaluates content validity by the thoroughness and care with which these operations have been conducted. In contrast, face validity is a judgment that the requirements of a test merely appear to be relevant. (p. 28)

Construct validity is implied when one evaluates a test or other set of operations in light of the specific construct;

> Judgments of construct validity are useful in efforts to improve measures for the scientific study of a construct. They are also useful when a test developer or test user wishes to learn more about the psychological qualities being measured by a test than can be learned from a single criterion-related validity coefficient.
>
> Evidence of construct validity is not found in a single study; rather, judgments of construct validity are based upon an accumulation of research results. In obtaining the information needed to establish construct validity, the investigator begins by formulating hypotheses about the characteristics of those who have high scores on the test in contrast to those who have low scores. Taken together, such hypotheses form at least a tentative theory about the nature of the construct the test is believed to be measuring. In a full investigation, the test may be the dependent variable in some studies and the independent variable in others. Some hypotheses may be "counterhypotheses" suggested by competing interpretations or theories.

> Such hypotheses or theoretical formulations lead to certain predictions about how people at different score levels on the test will behave on certain other tests or in certain defined situations. If the investigator's theory about what the test measures is essentially correct, most of his predictions should be confirmed. If they are not, he may revise his definition of the construct, or he may revise the test to make it a better measure of the construct he had in mind. Through the process of successive verification, modification, or elimination of hypotheses, the investigator increases his understanding of the qualities measured by the test. Through the process of confirmation or disconfirmation, test revision, and new research on the revised instrument, he improves the usefulness of the test as a measure of a construct. (p. 30)

Just as the proof of the reliability of instrumentation is required for the validation of a Title IV-C program, proof of the validity of the measures employed in a program evaluation must also be presented. Evidence inidcating the extent to which an instrument is measuring exactly what it is designed to measure is extremely difficult to gather. The process of validating an instrument becomes much more cumbersome than the determination of the reliability of an instrument, as one considers that different types of validity and the appropriate application of each. It is therefore mandatory that the director of a Title IV-C project be certain that the person responsible for the validity and reliability of the instrumentation

to be utilized in the summative evaluation of the innovation is qualified and responsible.

Data Analysis Techniques

Data analysis consists of organizing a quantity of data so that its meaning may be understood. Techniques of analyzing data range from a simple rank ordering of scores to very complex statistical treatment. Data analysis techniques allow the reader to identify relationships that are not apparent in the initial raw data and make it possible to compare a group or groups at different times (CSE, 1977, p. B5).

Effective use of statistical methods require careful distinction between formative evaluation and summative evaluation. "Failure to make the distinction between them has led to uninspired (use of) statistical methods and to misguided inferences" (Deming, 1975, p. 57). Simple descriptive statistics are appropriate for formative evaluation. Descriptive statistics describe data in terms of central tendency, variability and frequency. Measures of central tendency include the mode, median, and mean. Measures of variability include the standard deviation and range. Steiger, Fink and Kosecoff (1976), refer to descriptive statistics as "most useful in (Formative) evaluations because they are inherently meaningful and are the units for more

complex and less intuitive statistical procedures" (p. D2). Utilizing descriptive statistics in a formative evaluation provides continuous data about the development of a program on a day-to-day basis. This important formative feedback allows the decision maker to fine tune the innovation for future implementation and summative evaluation.

Sophisticated means of data analysis are necessary for summative evaluation where conclusive documentation of student change is necessary. Inferential statistics are appropriately included in summative evaluation designs. Inferential statistics provide a way to test the significance of results obtained when data are collected. In addition, inferential statistics provide a method to separate chance errors and random fluctuation from real change. Inferential statistics appropriate for inclusion in a summative evaluation design include:

Correlation - To measure the relationship between two variables.

Regression - To examine the relationship between a criterion (dependent) variable and two or more predictor (independent) variables.

Univariate Analysis of Variance (ANOVA) - Is a statistical technique used to compare two or more groups (independent variables) in terms of a single dependent variable.

Analysis of Covariance - Is a form of ANOVA in which the dependent variable is corrected or adjusted.

Multivariance Analysis of Variance (MANOVA) - Is analogous to ANOVA except that two or more dependent variables are analyzed together (Steiger, Fink and Kosecoff, 1976, p. D2).

Often, univariate statistical models as well as multivariate models are necessary in summative evaluation (Eber, 1975). Recent publications (Timm, 1975; Kolinger & Pedhazur, 1973; Sjoberg, 1975) make available new statistical tools for evaluators to use in summative data analysis.

When analyzing data with inferential statistics, it is important to make only valid inferences about the data. "Statistical inference (in a Summative Evaluation) is most effective when it is presented as conclusions valid for the frame studied and for the range of environmental conditions specified for the tests. It is important to make clear that conditions drawn by statistical theory may not hold under other conditions, and that other conditions may well be encountered" (Deming, 1975, p. 62).

In preparing the data analysis for the Summative Evaluation of a Title IV-C project, it is important to both utilize correct statistical procedures and make valid inferences from them.

Summary

The literature dealing with the evaluation of innovative education is extensive and suggests many

approaches to educational evaluation. Many good models for the different aspects of evaluation are presented in the literature. However, as in the past, the wealth of information is not easily accessible to Title IV-C project directors and evaluators. This important information should be catalogued in an evaluation resource for "hands-on" evaluation use. In addition, the evaluation resource should allow easy access to evaluate information pertinent to the developmental phase of the Title IV-C project. Sjoberg (1979) sums this up with "My intent, is not to rid social research of a standardized set of rules or ideal norms for collecting and analyzing data. My concern is with formulating these so that they can be more useful than those currently exposed" (p. 49).

The summative evaluation of a Title IV-C project should occur when a project is mature and students are affected by the project. The summative evaluation should be designed to meet the standards set forth by the Massachusetts Validation Process and the Joint Dissemination Review Panel (JDRP). The summative evaluation design should carefully include the appropriate use of research design, instrumentation and statistics. Instrumentation for the evaluation should be identified or developed before the summative evaluation is put in place. When a standardized norm-referenced test that

truly matches the program objectives does not exist, a more accurate criterion-referenced test must be developed. Summative evaluation is highly technical and should be conducted by an outside evaluation expert. However, the Title IV-C project director must retain control of the evaluation in order to guard against the evaluation shaping the project rather than the project shaping the evaluation.

C H A P T E R I I I

DESIGN OF THE STUDY

Introduction

Chapter III details the study design as well as the context with which this study was conducted. The data collection procedures are described in terms of instrumentation, development, categories of investigation, sample, data analysis and study relationship. Further, a model which depicts inter-relationships among data collection and data reporting is offered. The model also serves as a guide for the presentation of the results in Chapter IV.

Study Design

The study methodology was designed to ascertain why E.S.E.A. Title IV-C projects which are funded and expedited within the state of Massachusetts, routinely fail to measure up to validation guidelines established by the state and federal governments, and, to offer remedies, based upon data obtained, that are likely to resolve identified deficiencies. In order to reach this end, six data collection procedures, including the

Massachusetts Title IV-C Assessment, Massachusetts Validation Process, Archive Review, Interviews, Evaluation Design Review, and Assessment of Massachusetts Evaluation Needs, were utilized. The study design permitted the acquisition of data from every available source. Archive documentation was exhaustively reviewed, every individual who could provide information relevant to the study was questioned, and all available resources were utilized to employ personnel necessary to implement the Evaluation Design Review and the Interviews. These Data Collection Procedures provided information relevant to the characteristics of program evaluation including project director, program objectives, evaluation design, instrumentation, statistics, validation, evaluators, and Massachusetts Department of Education. Examination of the many categories of investigation enabled the investigator to document policies and practices and to identify problems which contribute to non-validation.

The design of the study described in Objective 1 is presented in Chapter III. The documentation of evaluation policies and practices and the identification of problems which contribute to the unsatisfactory evaluation outcomes (Objectives 1 and 2) are presented in Chapter IV. An evaluation plan designed to solve the evaluation problems identified within the Title IV-C program is presented in Chapter V.

Data Collection Procedures

Massachusetts Title IV-C Assessment

Description

Each year Title IV-C in Massachusetts is assessed on a Commonwealth-wide basis. The assessment is designed to monitor the progress of Title IV-C in Massachusetts, and to satisfy the specific data needs of the Massachusetts Title IV-C Coordinator and staff, Massachusetts Title IV Advisory Council, and the United States Office of Education. The investigator was one of three evaluators involved in the 1978 Assessment and the sole investigator of the 1979 Title IV-C Assessment.

Instrumentation

A questionnaire served as a primary source of the project data. The instrument was distributed to all Title IV-C project directors in Massachusetts for the purpose of obtaining desired perceptual data. The questionnaire was developed by the investigator according to the standards for questionnaire construction set forth by Kornhauser and Sheatsley (Selitiz, 1976). The same procedures were used for questionnaire development in both the 1978 and 1979 Assessments.

First, a Needs Analysis was conducted by the investigator, in order to formulate the precise

problem to be answered. The Title IV-C Coordinator, Title IV-C Staff, Massachusetts Title IV-C Advisory Council and Title IV-C Project Director were consulted and an assessment design was conceptualized. Next, the questionnaire topics were defined in an outline. Once in outline form, the ordering of topics was considered. The topics were examined for the best psychological sequence from the standpoint of the respondent. Each topic was analyzed for content and form, and put into questions. The number of questions necessary to gather sufficient data on each item was considered. These analyses were conducted by following the Kornhauser and Sheatsly checklist of points to consider in formulating questions titled "Guide for Questionnaire Construction" (p. 547). A few questions aimed at checking the reliability and consistency of responses were added to the questionnaire. Each of the resulting questionnaires were examined by the decision makers and revised.

The completed questionnaires were then pretested. Respondents reacted to the format and the appropriateness of the questions. Their responses were used to revise the instruments. Through this process the questionnaires were revised as shown in Appendix A.

Sample

The questionnaires were administered to Cycle 1 and Cycle II project directors. Cycle I project directors completed a questionnaire in the spring of their second year of project implementation (1978) and in the spring of the third year of project implementation (1979). Cycle II project directors completed a questionnaire in the spring of their first year of project implementation (1978) and in the spring of the second year of project implementation (1979). In 1978, twenty-four (24) Cycle I and thirty-six (36) Cycle II directors received and returned questionnaires. In 1979, twenty-four (24) Cycle I and thirty-six (36) Cycle II directors received questionnaires. Questionnaires were returned by twenty-one (21) Cycle I and thirty-five (35) Cycle II directors.

Categories of Investigation

The Massachusetts Title IV-C Assessment questionnaires include categories of investigation that examine the director's perception of project evaluation efforts. These categories of investigation are presented in Table 2 and include:

1. General project information
2. Knowledge of evaluation theory
3. Knowledge of evaluation application

Table 2

A Description of the Massachusetts Title IV-C Assessment

Data Collection Procedure	Modus Operandi			Categories of Investigation	Comment
	Instrumentation	Sample	Administration		
Massachusetts Title IV-C Assessment	Questionnaires	Project Directors	Cycle I: 1) spring of 2nd year of project implementation 2) spring of 3rd year of project implementation Cycle II: 1) spring of 1st year of project implementation 2) spring of 2nd year of project implementation	Knowledge of evaluation theory Knowledge of evaluation application Objectives -modification of project objectives for evaluation Evaluation -rating of project evaluation	The Project Director generally is responsible for the evaluation component of the project. It is important when studying project evaluation understand the views of the director concerning project evaluation

Table 2 (continued)

Data Collection Procedure	Modus Operandi			Categories of Investigation	Comment
	Instrumentation	Sample	Administration		
				-project components being evaluated ·student ·achievement ·student skill development ·student attitude ·personnel/teacher skill development ·cost effectiveness ·parent attitude ·community attitude Validation -directors' perceptions of Massachusetts Validation Process -directors' expected success at Massachusetts validation -directors' expected success at federal validation	

Table 2 (continued)

Data Collection Proecdure	Modus Operandi			Categories of Investigation	Comment
	Instrumentation	Sample	Administration		
				Project Evaluator -frequency of contact with evaluator -satisfaction of contact with evaluator -satisfaction with evaluators' level of understanding of project -help provided with evaluation, validation, program planning, program implementation Need for Training -Formative Evaluation -Summative Evaluation -Massachusetts Validation Requirements -Federal Validation Requirements Success with Massachusetts Validation	

4. Objectives

   - modification of project objectives for evaluation

5. Evaluation

   - rating of project evaluation

   - project components being evaluated

     · Student achievement
     · Student skill development
     · Student attitude
     · Personnel/teacher achievement
     · Personnel/teacher skill development
     · Personnel/teacher attitude
     · Cost effectiveness
     · Parent attitude
     · Community attitude

6. Validation

   - directors' perception of Massachusetts Validation Process

   - directors' expected success at federal validation

7. Project Evaluator

   - frequency of contact with evaluator

   - satisfaction of contact with evaluator

   - satisfaction with evaluators' level of understanding of project

   - help provided with evaluation, validation, program planning, program implementation

8. Need for Training

   - Formative Evaluation

   - Summative Evaluation

   - Massachusetts Validation Requirements

   - Federal Validation Requirements

9. Success with Massachusetts Validation

Data Analysis

The data were computer analyzed utilizing the Statistical Package for the Social Sciences (SPSS). Descriptive statistics were obtained as inferential statistics would not be appropriate. These data were presented as director perception.

Study Relationship

The project director is generally responsible for the evaluation component of the project. It is important when studying project evaluation to understand the views of the director concerning project evaluation. These data allowed the investigator to characterize the average project director as well as make judgments about the quality of project evaluation as perceived by the directors.

Massachusetts Validation Process

Description

When a Title IV-C project is nearing the end of the second year of implementation and positive summative evaluative data has been gathered, the director may apply for Massachusetts Validation. As described in Chapter I, the Massachusetts Validation Process is

designed to validate worthy innovations. The process is accomplished by an onsite visit by a validation team. The validation team consists of three members. At least one member of each team is considered to have expertise in evaluation, and is responsible for the team's judgments about the "effectiveness" of the project.

> Evidence of effectiveness is defined as the supporting evidence provided to show that the attainment of the major objective(s) can be attributed to the project activities. (MDE, 1979, p. 6)

The second team member is responsible for the team's judgments about the "exportability of the project. Exportability is determined from:

> Information provided to demonstrate that it is feasible to transport the program or practice to other school districts and that it can be adopted or adapted by other school districts. (p. 6)

The third team member is responsible for the team's judgments about the "cost effectiveness" of the project. Cost effectiveness is determined by a cost analysis. Cost analysis gathers:

> Information needed to describe the costs of start-up, operation and management, and the population to be served, which, then combined with evidence of effectiveness and exportability,

> will assist an interested shcool district to make an informed decision about adoption or adaption of the program or practice. (p. 6)

In addition, one of the team members is considered to be an expert in the content area of the project. During the onsite visit the validation team:

> (1) Reviews the objectives and activities of the project and discusses and shares scores on the three sections of the report; (2) observes a sample of project activities. (This could include a demonstration, audio-visual presentation, previsit, visiting a project site, etc.); (3) interviews personnel associated with the project as identified by the project director (advisory board, student, parents, teachers, administrators, etc.). The purpose of such interviews is to determine involvement, understanding and reaction to the project and to verify the fidelity of the data that have been provided; (4) convenes validation team to review validation materials and to develop the preliminary validation report for each of the three areas: <u>effectiveness</u>, <u>exportability</u>, and <u>cost analysis</u>; (5) holds an exit meeting with the project director and advises him/her of the validation team's recommendations; (6) forwards the validation team report, after completion by the chairperson, to the Massachusetts Department of Education, Title IV-C Coordinator. (MDE, 1979, p. 12)

Instrumentation

Following an on-site validation visit of a Title IV-C project, each validator completed a questionnaire. The questionnaire focused on data about program evaluation. The questionnaire was developed by the investigator according to the standards for questionnaire construction set by Kornhauser and Sheatsley (Selitz, 1976). First, an assessment design was conceptualized and the questionnaire topics were defined in an outline. Once in outline form, the ordering of topics was considered. The topics were examined for the best psychological sequence from the standpoint of the respondent. Each topic was analyzed for content and form, and put into questions. The number of questions necessary to gather sufficient data on each item was considered. The questionnaires were pretested. Respondents reacted to the format and the appropriateness of the questions. Through this process the questionnaires were revised as shown in Appendix B.

Sample

Following each on-site visit a questionnaire was completed by each of the validators who participated in the Massachusetts Validation Process. One hundred thirty-five (135) questionnaires were distributed to the validators serving the nineteen (19) Cycle I projects and twenty-six (26) Cycle II projects who applied for Massachusetts Validation. The validators returned ninety-nine (99) questionnaires.

Categories of Investigation

The Massachusetts Validation Process Questionnaire include categories of investigation that examine program evaluation. These are presented in Table 3 and include:

1. Evaluation Design
   - yield data that demonstrate change in the client group
   - yield data that demonstrate the project is responsible for change
2. Instrumentation
   - nationally normed
   - locally developed
3. Reliability
4. Validity
5. Statistics

Table 3

A Description of the Massachusetts Validation Process Questionnaire

Data Collection Procedure	Modus Operandi			Categories of Investigation	Comments
	Instrumentation	Sample	Administration		
Massachusetts Validation Process Questionnaire	Questionnaire	Validators	Each validator completed a questionnaire following an on-site visit to a Title IV-C project	Evaluation Design -yield data that demonstrate change in the client group -yield data that demonstrate the project is responsible for change Instrumentation -nationally normed -locally developed Reliability Validity Statistics	A 3 member validation team conducts an on-site visit to each project whose application is accepted. At least one team member is an evaluation expert

Data Analysis

The data were computer analyzed utilizing the Statistical Package for the Social Sciences (SPSS). Descriptive Statistics were employed, the use of inferential statistics would be inappropriate.

Study Relationship

These data allowed the investigator to make judgments about the appropriate use of project evaluation for validation as reported by trained validators.

Archive Review

Description

An archive review was conducted. The documents reviewed were the Massachusetts Validation Team Reports. The Massachusetts Validation Team Reports are documents completed by the Validation Team members following each on-site visit. The reports include the Team's judgments about the effectiveness of the project as well as comments and suggestions.

Categories of Investigation

The investigator collected data on the category of investigation that deals with project evaluation efforts. This category is presented in Table 4:

Table 4

A Description of the Archive Review and Interview Process

Data Collection Procedure	Modus Operandi			Categories of Investigation
	Instrumentation	Sample	Administration	
Archive Review	Massachusetts Validation Team Reports		Validation Report compiled by validation team after an on-site visit	Success at Massachusetts Validation
Interviews	Interview	Title IV-C Coordinator		Perceived quality of evaluation
	Interview	Title IV-C Staff		

1. Success with Massachusetts Validation

Reviewing these documents provided data concerning the rate of project success with the Massachusetts Validation Process.

Study Relationship

Utilizing these data, the investigator made judgments about the appropriate use of evaluation for validation by projects that were validated and by projects that were not validated.

Interviews

Description

Interviews were conducted with a sample of the Massachusetts Department of Education Title IV-C Administrators. The interviews were conducted by an evaluator other than the study investigator. The interviews emphasized the acquisition of data about opinions and views of Massachusetts Title IV-C personnel on the perceived quality of program evaluation.

Instrumentation

An interview questionnaire was developed according to the standard for questionnaire construction set by Kornhauser and Sheatsley (Sleitz, 1976). First, the

questionnaire topics were defined in an outline. Once in outline form, the ordering of topics was considered. Each topic was analyzed for content and form, and organized in the form of questions. The questionnaire was pretested. Respondents reacted to the format and the appropriateness of the questions. Through this process the questionnaire was revised as shown in Appendix C. The questionnaires were designed to gather data on several categories of investigation relevant to the implementation of Title IV-C. One of these categories examines program evaluation and is relevant to this study.

The interviews were conducted by an evaluator other than the study investigator. The use of an external evaluator provided assurance that the opinions of the investigator were not projected on to those being interviewed. The interviews were conducted in accordance with the standards for interviewing set by Kornhauser and Sheatsley (Selitz, 1976).

Categories of Investigation

The interviews focused upon the opinions and views of Massachusetts Title IV-C personnel on one category of investigation, the perceived quality of program evaluation. This information is presented in Table 4.

Sample

The Massachusetts Title IV-C Coordinator and five (5) of the six (6) members of the Title IV-C staff were interviewed (N=6). The sixth staff member was out of the country during the interview period.

Study Relationship

These data were reported in support of the need to improve the evaluation of Title IV-C projects in Massachusetts.

Evaluation Design Review

Description

Independent evaluators who have had experience with the JDRP and who are considered to be experts in educational evaluation reviewed Title IV-C program evaluation designs for worthiness. The evaluators who participated in the review were Louis Aikman, Ph.D., Boston University; Ronald Nuttal, Ph.D., Boston College; David J. Rosen, Ed.D., University of Massachusetts; Gene Mulcahy, Ed.D., Hartford Public Schools; Robert Algozzine, University of Florida; and Martin J. Higgins, Ph.D., West Chester State College.

Categories of Investigation

The evaluation design review focused upon data on six categories of investigation. They are presented in Table 5 and include:

1. Objectives
2. Evaluation Design
3. Instrumentation
4. Reliability and Validity
5. Sample
6. Statistics
7. Reporting

Sample

The sample included thirty-two (32) Title IV-C projects that received initial funding in 1977. The Evaluation Design Review was made available to the thirty-seven (37) directors of Title IV-C projects then, in their first year of project implementation. The sample includes all the directors who wanted the service.

Study Relationship

This critical analysis allowed the investigator to present an unbiased external judgment about the quality of Title IV-C evaluation in Massachusetts.

Table 5

A Description of the Evaluation Design Review

Data Collection Procedure	Modus Operandi			Variables
	Instrumentation	Sample	Administration	
Evaluation Design Review	Evaluation Designs	Cycle II Projects Directors		Evaluations will be reviewed for: Objectives Evaluation Design Instrumentation Reliability and Validity Sample Statistics Reporting

Assessment of Massachusetts Evaluation Needs

Description

A conference was held to generate a comprehensive list of the evaluative problems of Title IV-C projects. The requirements for validation set by the Massachusetts Validation Process and the Joint Dissemination Review Panel (JDRP) were also discussed. These data reflect the degree to which Massachusetts Title IV-C projects have attained the prescribed evaluation standards. Participants included United States Office of Education staff, Massachusetts Department of Education staff, National Diffusion Network staff, validators, project directors, evaluation experts, and the study investigator.

Categories of Investigation

Five categories were addressed at the conference. They include:

1. Evaluation Resources
2. Objectives
3. Federal Validation
4. Evaluators
5. Instrumentation

These categories are presented in Table 6.

Table 6

A Description of the Assessment of Massachusetts Evaluation Needs

Data Collection Procedure	Modus Operandi			Variables
	Instrumentation	Sample	Administration	
Assessment of Massachusetts Evaluation Needs	Conference/ Expert Testimony	Participants include USOE, MDE, NDN, Validators, Project Directors, evaluation experts		Evaluation resources Objectives Federal Validation Evaluators Instrumentation

Study Relationship

These data will allow the investigator to compare Massachusetts Title IV-C evaluation efforts to federal standards.

Summary

Six data collection procedures including the Massachusetts Title IV-C Assessment, Massachusetts Validation Process, Archive Review, Interviews, Evaluation Design Review, and Assessment of Massachusetts Validation Needs, were utilized in order to study the evaluation procedures of Title IV-C in Massachusetts. These data were gathered in order to systematically study the state of Massachusetts' Title IV-C Operation and to document evaluation policies and practices.

Data from the six collection procedures were analyzed and synthesized, and conclusions were drawn. These data provided a picture of project evaluations as they are currently implemented in Massachusetts. This information was grouped to point out specific weaknesses in formative evaluation and summative evaluation, with an emphasis on instrumentation, research design, statistics, and success at validation. Specific evaluation inadequacies were identified and reported as part of Chapter IV.

An evaluation plan — based upon data reported — that will be implemented according to the development phase of the project is proposed in Chapter V. The evaluation plan is currently being implemented in Massachusetts.

A study model was developed by the investigator in order to guide the presentation of the results in Chapter IV. Table 7 delineates the Categories of Investigation as well as the methodology utilized by the investigator to collect and present results in the following chapter of the study.

Table 7

Study Model for Guiding the Presentation of the Results

Categories of Investigation	Descriptors	Method of Investigation
Project Director	Background information	Massachusetts Title IV-C Assessment
	Perception of program evaluation	
	Training needed in program evaluation	
Program Objectives	Address key components of project	Evaluation Design Review
	Completeness	
	Operational	
	Objectives modified	Massachusetts Title IV-C Assessment
Evaluation Design	Evaluation design demonstrates project is responsible for change	Massachusetts Validation Process Questionnaire
	Change in client group demonstrated	
	Nature of evaluation -formative -summative	Evaluation Design Review

Table 7 (continued)

Categories of Investigation	Descriptors	Method of Investigation
Evaluation Design (continued)	Growth due to project	Evaluation Design Review
	Treatment vs. non-treatment comparison	
	Quasi-experimental design	
	Treatment group similar to control group	
	Repetition in future in other schools	
	Data collection schedule	
	Evaluation personnel	
	Evaluation resources	Assessment of Massachusetts Evaluation Needs
	Validation focus	
	Skill level of evaluation personnel	
	Outcomes of evaluation design	
Statistics	Employed properly	Massachusetts Validation Process Questionnaire
	Statistics utilized	Evaluation Design Review
	Utilization clear	

Table 7 (continued)

Categories of Investigation	Descriptors	Method of Investigation
Validation	Director belief that project will be validated	Massachusetts Title IV-C Assessment
	N projects evaluating for presentation to JDRP	
	Director perception of Massachusetts Validation Process	
	Success with Massachusetts Validation	
Evaluators	Technical assistance provided by evaluators	Massachusetts Title IV-C Assessment
	Evaluation components	
Massachusetts Department of Education	Need for (1) statewide evaluation improvement, (2) evaluation management system, (3) evaluation support system	Interviews
	Reporting schedules and procedures	Evaluation Design Review
	Report audiences	
Instrumentation	Nationally normed instruments, locally developed instruments	Massachusetts Validation Process Questionnaire
	Instrument development procedures	

Table 7 (continued)

Categories of Investigation	Descriptors	Method of Investigation
Instrumentation (continued)	Reliability	
	Validity	
	Measures "fit" objectives	Evaluation Design Review
	Test administration procedures explained/acceptable	
	Procedures for training interviewers/observers/ administrators explained/ acceptable	
	Sample defined	
	Sample size	
	Sampling procedures	

C H A P T E R I V

PRESENTATION AND ANALYSIS OF THE RESULTS

The chapter presents and analyzes data collected in support of the three study objectives. The study objectives are:

Objective 1: To systematically study the state of Massachusetts' Title IV-C operation in order to document evaluation policies and practices.

Objective 2: To identify problems, based upon data obtained via surveys and document analyses, which contribute to the program's unsatisfactory evaluation outcomes.

Objective 3: To offer an evaluation plan which is tailored to resolve evaluation problems identified within the state's Title IV-C program.

The model presented in Chapter III of the study is used as the format for the presentation of data. Chapter IV provides a presentation and analysis of the data for each of the categories of investigation. The categories include, the Project Director, Program Objectives, Evaluation Design, Instrumentation, Statistics, Validation, Evaluators, Massachusetts Deparment of Eudcation.

The following data presentation will document evaluation policies and practices as well as identify problems which contribute to the program's unsatisfactory evaluation outcomes.

Presentation of the Results

Objective 1

The Project Director

Several questionnaire items examined the background of the Cycle I and Cycle II project directors. These data are presented in Table 8. The directors were asked, "How did you come to be director of your project?" It can be seen that over half of the project directors (55%) wrote the proposal for their projects, and an additional thirty-two percent (31.7%) helped plan their projects. Thirty-five percent (35%) were hired from outside the system and twenty-five (25.0%) were hired from within the system.

The director's employment status immediately prior to appointment as director was examined. Nearly fifty-two percent (51.6%) of the project directors were teachers immediately before becoming project directors. Nineteen percent (19.3%) acted as department chairpersons and nearly thirteen percent (12.9%) were school building administrators. Ten percent (9.6%) were guidance counselors and seven percent (6.4%) were school district administrators.

Table 8

Results of the Project Directors' Responses to Questionnaire Items Focused on Their Background

Item	Type of Response	Number and Percent of Responses: Number	Percent
How did you come to be director of your project?	Helped plan project	19	31.7
	Wrote proposal	33	55.0
	Hired from within system	15	25.0
	Hired from outside the system	25	35.0
	Other	1	1.7
What was your position immediately prior to your appointment as director?	Teacher	16	51.6
	Department Chairperson	6	19.3
	Guidance Counselor	3	9.6
	School Building Administrator	4	12.9
	School District Administrator	2	6.4
How many years have you been in your school district?	Less than 1 year	8	14.3
	1-5 years	28	50.0
	5-10 years	12	21.4
	Over 10 years	8	14.3
What is the highest degree you have earned?	Bachelors Degree	17	28.3
	Masters Degree	31	51.7
	C.A.G.S.	8	13.3
	Doctoral Degree	4	6.7
How would you rate the evaluation componnet of your project?	1978		
	Excellent	5	13.9
	Very Good	10	27.8
	Good	16	44.4
	Poor	4	11.1
	Very Poor	1	2.8
	1979		
	Excellent	7	21.2
	Very Good	9	27.3
	Good	10	30.3
	Poor	6	18.2
	Very Poor	1	3.0

The directors were asked to indicate the number of years they were employed by their school districts. Half of the directors (50.0%) were in the school system between 1 and 5 years. Over twenty-one percent (21.4%) were in the school system from 5 to 10 years. And over twenty-eight percent (28.6%) were in the school system for less than 1 year or over 10 years.

The directors' education was examined. Over half of the directors (51.7%) have Masters Degrees, thirteen percent (13.3%) have Certificates of Advanced Graduate Study, and nearly seven percent (6.7%) have doctorates.

The 1978 and 1979 Massachusetts Title IV-C Assessment Questionnaires examined the Cycle II director's perception of Project Evaluation.

In 1979, twenty-one percent (21.2%) of the Cycle II project directors felt they had excellent evaluation designs, twenty-seven percent (27.3%) very good evaluation designs, thirty percent (30.3%) good, eighteen percent (18.2%) poor, three percent (3.0%) indicated very poor. In 1978, fourteen percent (13.9%) of the Cycle II directors indicated that their evaluation designs were excellent. Further, twenty-eight percent (27.8%) indicated very good, forty-four percent (44.4%) good, eleven percent (11.1%) poor, three percent (2.8%) indicated very poor.

The Massachusetts Title IV-C Assessment included items that examined the evaluation training needed by project directors. These data are presented in Table 9.

The directors of both Cycle I and Cycle II projects reported a need for training in formative evaluation (28.6%, Cycle I; 29.4%, Cycle II), summative evaluation (38.1%, Cycle I; 32.4%, Cycle II), data utilization (23.8%, Cycle I; 55.9%, Cycle II), Massachusetts Validation Requirements (4.8%, Cycle I; 11.8%, Cycle II), and Federal Validation Requirements (23.8%, Cycle I, 50.0%, Cycle II). In addition, some of these directors reported no need for training in formative evaluation (14.3%, Cycle I; 29.4%, Cycle II), summative evaluation (4.8%, Cylce I; 29.4%, Cycle II), data utilization (14.3%, Cycle I; 26.5%, Cycle II), Massachusetts Validation Requirements (28.6%, Cycle I; 26.5%, Cycle II), and Federal Validation Requirements (19.0%, Cycle I; 32.4%, Cycle II).

Program Objectives

The Evaluation Design Review produced data relevant to program objectives. These data are presented in Table 10.

Table 9

The Project Directors' Response to the Question "In which of the following areas do you feel you need training?"

Item	Percent of Responses in each category			
	Cycle I Project Directors*		Cycle II Project Directors	
	No Need for Training	Need Training	No Need for Training	Need Training
Formative Evaluation	14.3	28.6	29.4	29.4
Summative Evaluation	4.8	38.1	29.4	32.4
Data Utilization	14.3	23.8	26.5	55.9
Massachusetts Validation Requirements	28.6	4.8	26.5	11.8
Federal Validation Requirements	19.0	23.8	32.4	50.0

*The directors of validated Cycle I projects.

Table 10

Results of the Evaluation Design Review and the Massachusetts Title IV-C Assessment Relevant to Program Objectives

Item	Type of Response	Number and Percent of Responses	
		Number	Percent
Program Objectives: Address key components of project	Yes	23	71.9
	No	9	28.1
Completeness	Yes	9	28.1
	No	20	62.5
	Evaluation Design not sufficiently sensitive to provide such information	2	6.3
	Some, but not all objectives	1	3.1
Operational (Observational & Measurable)	Yes	6	18.8
	No	23	71.9
	Evaluation Design not sufficiently sensitive to provide such information	1	3.1
	Some, but not all objectives	2	6.3

Seventy-two percent (71.9%) of the program evaluations addressed the key components of the projects, twenty-eight percent (28.1%) did not.

The reviewers indicated that twenty-eight percent (28.1%) of the evaluation designs included objectives that completely addressed all the key outcomes of the project, sixty-three percent (62.5%) did not. Further, the reviewers indicated that six percent (6.3%) of the evaluation designs were not sufficiently sensitive to provide such information. Some, but not all of the objectives were complete for three percent (3.1%) of the projects.

Data indicates that nineteen percent (18.8%) of the program evaluation designs included objectives that were both observable and measurable, seventy-two percent (71.9%) did not. Three percent (3.1%) of the evaluation designs were not sufficiently sensitive to provide this information. Further, six percent (6.3%) of the evaluation designs included some (but not all) objectives in operational form.

The Massachusetts Title IV-C Assessment Questionnaires included an item that examined program objectives. These data are presented in Table 11.

Seventy-eight percent (78.4%) of the Cycle II project directors reportedly modified their program objectives in their first year of project implementation

Table 11

The Project Directors' Response to the Question "Have your project objectives been modified since your project began?"

	Number and Percent of Responses	
	Cycle II Project Directors	
Type of Response	Number	Percent
1979		
Yes	29	85.3
No	5	14.7
1978		
Yes	29	78.4
No	8	21.6

(1978) and twenty-two percent (21.6%) did not. Eighty-five percent (85.3%) of these directors reportedly modified their program objectives their second year of project implementation (1979) and fifteen percent (14.7%) did not.

Evaluation Design

The Massachusetts Validation Process Questionnaire included items which examined program evaluation design. These data are presented in Table 12.

In Table 12 are the validators response to the question "Did the evaluation design for this project yield data that demonstrates that the project is responsible for the change?" Most (77.6%) of the validators felt the evaluation demonstrated that the project was responsible for change in the client group. Over twenty-two percent (22.4%) of the validators indicated that the evaluation design did not demonstrate that the project was responsible for change.

In Table 12 are the validators' responses to the question "Did the evaluation design for this project yield data that demonstrate change (growth) in the client group?" Over eighty-three percent (83.6%) of the validators indicated that projects validated for diffusion had evaluation designs that demonstrated change in the client group. Eighty-two percent (82.0%) of the validators indicated

Table 12

Results of the Validators Response to Questionnaire Items Focused on Program Evaluation Design

Item	Type of Response	Number and Percent of Responses	
		Number	Percent
Did the evaluation design for this project yield data that demonstrate that the project is responsible for the change?	Yes	52	77.6
	No	15	22.4
Did the evaluation design for this project yield data that demonstrate change (growth) in the client group?	Projects Validated for Diffusion		
	Yes	51	83.6
	No	10	16.4
	Projects Not Validated for Diffusion		
	Yes	7	18.0
	No	31	82.0

that the projects not validated for diffusion had evaluation designs that did not yield data that demonstrated change in the client group.

The evaluation experts who participated in the Evaluation Design Review, examined the evaluation designs of the Cycle II projects. These data are presented in Table 13.

Data indicates that nearly twenty percent (18.8%) of the Cycle II directors had evaluation designs that were only formative in nature. Eighty-one percent (81.3%) of these projects had evaluators who designed a more sophisitcated evaluation.

Of the Cycle II projects, sixteen percent (15.6%) had evaluation designs that permited the experts to conclude that the observed growth (change) of the project was indeed due to the project. Eighty-four percent (84.4%) did not.

Twenty-nine percent (28.1%) of the evaluation designs included a comparison of treatment vs. nontreatment conditions. Nearly seventy-two percent (71.9%) did not.

The experts indicated that six percent (6.3%) of the evaluators utilized a quasi-experimental design. Ninety-three percent (93.8%) did not.

Table 13

Results of the Evaluation Design Review Relevant to Evaluation Design

Item	Number and Percent of Responses					
	Response					
	Yes		No		Total	
Evaluation Design:						
Formative Evaluation only	6	(18.8)	26	(81.3)	32	(100.0)
Observed growth due to project	5	(15.6)	27	(84.4)	32	(100.0)
Treatment vs. non-treatment comparison	9	(28.1)	23	(71.9)	32	(100.0)
Quasi-experimental design utilitized	2	(6.3)	30	(93.8)	32	(100.0)
Treatment group is similar to control group	1	(3.1)	31	(96.9)	32	(100.0)
Future repetition in other schools	3	(9.4)	29	(90.6)	32	(100.0)
Data collection schedule included	4	(12.5)	28	(87.5)	32	(100.0)
Evaluation personnel listed	1	(3.1)	31	(96.9)	32	(100.0)

Three percent (3.1%) of the evaluations included documentation that the treatment group is similar to the control group. Further, ninety-seven percent (96.9%) did not include this documentation.

Nine percent (9.4%) of the evaluators included provisions for repetition of the design in adopter schools in the future. Ninety percent (90.6%) did not.

Thirteen percent (12.5%) of the project evaluators included data collection schedules in their evaluation designs, eighty-eight percent (87.5%) did not.

Three percent (3.1%) of the evaluators included a list of the personnel responsible for each job, ninety-seven percent (96.9%) did not.

The participants of the conference to Assess Massachusetts Evaluation Needs identified characteristics of project evaluation designs and made recommendations.

Five major characteristics of project evaluation designs emerged from the conference: (1) projects were under-resourced in funds to carry out adequate evaluation designs; (2) project goals were too broad in scope; (3) evaluation designs were focused on the Massachusetts Validation Process rather than on the goal of validation for national dissemination by the Joint Dissemination Review Panel; (4) implementation of the evaluation design

relied, in most instance, on unskilled persons; (5) evaluation designs focused on small ends rather than on outcomes that will be educationally significant. Further, it was recommended that guidelines be developed, which would convey to project directors and project evaluators that evaluation designs for Title IV-C projects should be of such a quality that it would enhance a project's capability of being initially validated for diffusion by the Massachusetts Validation Process and subsequently validated for national dissemination by the Joint Dissemination Review Panel.

The evaluation experts who participated in the Evaluation Design Review examined the reporting procedures described by the evaluators of Cycle II projects in their evaluation designs. These data are presented in Table 14.

The evaluation experts indicated that nine percent (9.4%) of the program evaluators included a reporting schedule in their evaluation designs, nineteen percent (18.8%) did not. A further review of the data, indicates that six percent (6.3%) of the evaluators included a definition of the reporting procedures, nineteen percent (18.8%) did not. In addition, six percent (6.3%) of the evaluators identified each of the audiences he/she would report evaluative information to, twenty-two percent (21.9%) did not.

Table 14

Results of the Evaluation Design Review Relevant to Reporting

Item	Number and Percent of Response: Response: Yes No.	Yes %	No No.	No %	Information Not Provided By Review No.	Information Not Provided By Review %	Total No.	Total %
Reporting:								
Schedule	3	9.4	6	18.8	23	71.9	32	100.0
Included	3	33.3	6	66.7	--	--	9	100.0
Reporting Procedures	2	6.3	6	18.8	24	75.0	32	100.0
Included	2	25.0	6	75.0	--	--	8	100.0
Audiences	2	6.3	7	21.9	23	71.9	32	100.0
Defined	2	22.2	7	77.8	--	--	9	100.0

Instrumentation

The Massachusetts Validation Process Questionnaire included items that examined the instrumentation utilized in program evaluation. These data are presented in Table 15 and Table 16.

In Table 15 are the validator's responses to the question "The data obtained for the evaluation of this project were obtained from: Nationally Normed Instruments; Locally Developed Instruments; Both?"

Data indicates strong similarities in the types of instrumentation utilized by both validated and not validated projects. The majority (69.3%) of the evaluators utilized locally developed instruments. Nearly twenty-three percent (22.7%) of the evaluators used both Nationally Normed Instruments and Locally Developed Instruments, eight percent (8.0%) used only Nationally Normed Instruments. Nearly fifteen percent (14.8%) of the projects that were not validated for diffusion obtained evaluative data from Nationally Normed Instrumentation.

The validators responded to the question "If locally developed instruments were used to collect data, were the procedures used to develop the instruments explained?" The majority (77.2%) of the evaluators of projects that were validated for diffusion and used locally developed instruments to collect data explained

Table 15

The Validators' Response to Questionnaire Items Relevant to Instrumentation

Item	Type of Response	Number and Percent of Responses: All Projects		Projects Validated for Diffusion		Projects Not Validated for Diffusion	
The data obtained for the evaluation of this project were obtained from: Nationally Normed Instruments, Locally Developed Instruments; both?	Nationally Normed Instruments	7	8.0	3	4.9	4	14.8
	Locally Developed Instruments	61	69.3	42	68.9	19	70.4
	Both	20	22.7	16	26.2	4	14.8
If locally developed instruments were used to collect data, were the procedures used to develop the instruments explained?	Yes	--	--	44	77.2	10	35.8
	No	--	--	13	22.8	16	57.1
	Evaluation Design did not include the use of Instrumentation	--	--	0	--	2	7.1

Table 15 (continued)

Item	Type of Response	Number and Percent of Responses: All Projects		Projects Validated for Diffusion		Projects Not Validated for Diffusion	
Are the instruments reliable?	Yes	--	--	18	31.1	3	10.0
	No	--	--	5	8.6	7	23.3
	Information not included in Evaluation Design	--	--	35	60.3	20	66.7
Are the instruments valid?	Yes	--	--	6	10.5	1	3.4
	No	--	--	23	40.4	8	27.7
	Information not included in Evaluation Design	--	--	28	49.1	20	68.9

the procedures used to develop the instruments, over thirty-five percent (35.8%) of the projects not validated did not. More than half (57.1%) of the evaluators of projects that were not validated for diffusion and used locally developed instruments to collect data, did not explain the procedures used to develop the instruments.

The validators indicated that the majority of the evaluation designs for both projects validated for diffusion (60.3%) and projects not validated for diffusion (66.7%) did not include information on instrument reliability. Over thirty-one percent (31.1%) of the projects validated for diffusion used reliable instruments, nearly nine percent (8.6%) did not. Nearly a quarter (23.3%) of the projects not validated for diffusion did not use reliable instrumentation, ten percent (10.0%) did.

In Table 15 are the validators' responses to the question "Are the instruments valid?" The validators indicated that the majority of the evaluation designs for projects validated for diffusion (49.1%) and projects not validated for diffusion (68.9%) did not include information on instrument validity. Over forty percent (40.4%) of the projects validated for diffusion did not use valid instrumentation, nearly eleven percent (10.5%)

did. Nearly twenty-eight percent (27.7%) of the projects not validated for diffusion did not use valid instrumentation while approximately three percent (3.4%) did.

The evaluation experts who participated in the Evaluation Design Review, examined the Cycle II project evaluators' utilization of instrumentation. These data are presented in Table 16.

The evaluation experts indicated that three percent (3.1%) of the project evaluators included test development procedures in their designs, seventy-two percent (71.9%) did not.

Of the Cycle II projects, only three percent (3.1%) of the evaluators included information concerning the reliability and validity of the instrumentation utilized in the evaluation. Over ninety-three percent (93.8%) did not include reliability and validity information.

Data indicates that sixteen percent (15.6%) of the project evaluators choose measures that appropriately "fit" the program objectives, six percent (6.3%) chose measures that were appropriate for some, but not all the objectives. Further, seventy-eight percent (78.1%) of of the evaluators did not chose measures that appropriately "fit" the program objectives.

Table 16

Results of the Evaluation Design Review
Relevant to Instrumentation

Item	Number and Percent of Response									
	Response									
	Yes		No		Some, but not all Objectives		Information not Provided by Reviews		Total	
	No.	%	No.	%	No.	%	No.	%	No.	%
Instrumentation:										
Test development included	1	3.1	23	71.9	--	--	8	25.0	32	100.0
Test reliability included	1	3.1	30	93.8	--	--	1	3.1	32	100.0
Test validity included	1	3.1	30	93.8	--	--	1	3.1	32	100.0
Measures "fit" objectives	5	15.6	25	78.1	2	6.3	--	--	32	100.0
Test administration procedures explained	2	6.3	10	31.3	--	--	20	62.5	32	100.0
	2	16.7	10	83.3	--	--	--	--	12	100.0
Test administration procedures acceptable	1	3.1	1	3.1	--	--	30	93.8	32	100.0
Procedures for training Interviewers/Observers/ Test Administrator explained	2	6.3	10	31.3	--	--	20	62.5	32	100.0
	2	16.7	10	83.3	--	--	--	--	12	100.0
Procedures for training Interviewers/Observers/ Test Administrator Acceptable	1	3.1	1	3.1	--	--	30	93.8	32	100.0

Six percent (6.3%) of the project evaluators included an explanation of the procedures used to administer the testing program for the evaluation, thirty-one percent (31.3%) did not. Further, the evaluation experts found that three percent (3.1%) of the evaluation designs included acceptable test administration procedures.

The evaluation experts indicated that six percent (6.3%) of the evaluation designs included an explanation of the procedures utilized for training interviewers, observers, and test administrators, thirty-one percent (31.3%) did not. Three percent (3.1%) of the evaluation designs that included procedures for training interviewers, observers and test administrators were deemed acceptable by the experts.

The evaluation experts who participated in the evaluation design review examined the sampling procedures utilized by program evaluators. These data are presented in Table 17.

Fifty percent (50.0%) of the evaluators included a definition of the evaluation sample in their evaluation designs, twenty-two percent (21.9%) did not.

The evaluation experts indicated that sixteen percent (15.6%) of the evaluators chose a correct sample size, thirty-one percent (31.3%) did not.

Table 17

Results of the Evaluation Design Review Relevant to Sampling

Item	Number and Percent of Response Response Yes No.	%	No No.	%	Information Not Provided By Reviewers No.	%	Total No.	%
Defined	16	50.0	7	21.9	9	28.1	32	100.0
	16	69.6	7	30.4	--	--	23	100.0
Correct Size	5	15.6	10	31.3	17	53.1	32	100.0
	5	33.3	10	66.7	--	--	15	100.0
Sampling Procedures Explained	9	28.1	7	21.9	16	43.8	32	100.0
	9	50.0	7	50.0	--	--	16	100.0
Sampling Procedures Acceptable	5	15.6	4	12.5	23	71.9	32	100.0
	5	55.5	4	44.5	--	--	9	100.0

Data indicates that twenty-eight percent (28.1%) of the evaluators included a description of their sampling procedures while twenty-two percent (21.9%) did not. The evaluation experts found sixteen percent (15.6%) of the procedures used for sampling to be acceptable, thirteen percent (12.5%) were not.

Statistics

The Massachusetts Validation Process Questionnaire included items that examined the validators' views of the use of statistics in program evaluation. These data are presented in Table 18.

Of the projects that were validated for diffusion, over sixty-four percent (64.3%) of the program evaluators properly used statistical procedures to analyze data, approximately seven percent (7.1%) did not. Further, nearly twenty-nine percent (28.6%) of these evaluation designs did not include the use of statistics. Half (50.0%) of the evaluators of projects not validated for diffusion did not include the use of statistics in program evaluation designs. Thirty-six percent (35.7%) of these directors did not utilize statistical procedures properly while over fourteen percent (14.3%) did.

The evaluation experts who participated in the evaluation review, examined the inclusion of statistical

Table 18

The Validators' Response to the Question "If statistical procedures were used to analyze/compare data, were they properly employed?"

Response	Number and Percent of Response			
	Projects Validated for Diffusion		Projects Not Validated for Diffusion	
	No.	%	No.	%
Yes	36	64.3	4	14.3
No	4	7.1	10	35.7
Evaluation Design did not include the use of statistics	16	28.6	14	50.0
TOTAL	56	100.0	28	100.0

procedures in program evaluation. These data are presented in Table 19.

The evaluation experts indicated that twenty-five percent (25.0%) of the evaluators included the use of statistical procedures in their evaluation designs, thirty-four percent (34.4%) did not. Forty-one percent (40.6%) of the reviewers did not provide this information.

Data indicates that nine percent (9.4%) of the program evaluators included a clear description of the statistical procedures they intended to utilize in the evaluation, sixteen percent (15.6%) did not.

Information concerning which analysis techniques the evaluator intends to use to analyze which data was included in nine percent (9.4%) of the designs, sixteen percent did not include this information.

Validation

The Massachusetts Title IV-C Assessment Questionnaires included two Categories of Investigation that examined Federal Validation. These data are presented in Table 20.

The Cycle II project directors were asked the question "To what extent do you believe that the project will be validated for Federal Dissemination?" Five percent (5.0%) strongly agree that their projects will

Table 19

Results of the Evaluation Design Review Relevant to Statistics

Item	Number and Percent of Response							
	Response							
	Yes		No		Information Not Provided By Reviewers		Total	
	No.	%	No.	%	No.	%	No.	%
Statistics:								
Utilization of	8	25.0	11	34.4	13	40.6	32	100.0
statistical procedures	8	42.1	11	57.9	--	--	19	100.0
Utilization	3	9.4	5	15.6	24	75.0	32	100.0
clear	3	37.5	5	62.5	--	--	8	100.0
Utilization of analysis	3	9.4	5	15.6	24	75.0	32	100.0
techniques clear	3	37.5	5	62.5	--	--	8	100.0

be validated for Federal Dissemination, twenty-five percent (25.0%) agreed, fifty percent (50.0%) were neutral, fifteen percent (15.0%) disagreed, five percent (5.0%) strongly disagreed.

The Cycle II project directors were asked the question "Is your project being evaluated for presentation to the Joint Dissemination Review Panel?" A review of the data indicates that seventeen percent (16.7%) of the Cycle I projects that were validated for diffusion are being evaluated for presentation to the Joint Dissemination Review Panel and eighty-three percent (83.3%) are not.

The Massachusetts Title IV-C Assessment included Categories of Investigation that examined the directors' views of the Massachusetts Validation Process. In 1978 and 1979 the directors of Cycle II projects viewed the Massachusetts Validation Process as necessary (25.7%, 1978; 29.4%, 1979) and an opportunity (45.7%, 1978; 35.2%, 1979). Further, twenty-six percent (26.4%) of the directors found the process to be a hindrance in 1979. In 1978, only six percent (5.7%) of the directors felt validation was a hindrance.

The Massachusetts Title IV-C Assessment Questionnaire included an item that examined the Cycle II project directors' hopes of being validated for diffusion. Data indicates that forty-seven percent (46.5%) of the Cycle II

Table 20

Results of the Project Directors' Responses to Questionnaire Items Focused on Validation

Item	Type of Response	Number and Percent of Responses: Number	Percent
To what extent do you believe that the project will be validated for Federal Dissemination (pass JDRP)?+	Strongly Agree	1	5.0
	Agree	5	25.0
	Neutral	10	50.0
	Disagree	3	15.0
	Strongly Disagree	1	5.0
Is your project being evaluated for presentation to the Joint Dissemination Review Panel?*	Yes	2	16.7
	No	10	83.3
The following describes my feeling about the Massachusetts Validation Process:+	1978		
	Helpful	--	17.1
	Hindrance	--	5.7
	Necessary	--	25.7
	Unnecessary	--	5.7
	An Opportunity	--	45.7
	No Feeling	--	25.7
	1979		
	Helpful	--	11.7
	Hinderance	--	26.4
	Necessary	--	29.4
	Unnecessary	--	14.7
	An Opportunity	--	35.2
	No Feeling	--	8.8

Table 20 (continued)

Item	Type of Response	Number and Percent of Responses: Number	Percent
To what extent do you believe that the project will be validated for Massachusetts Diffusion?+	Strongly Agree	13	46.5
	Agree	12	42.8
	Neutral	3	10.7
	Disagree	--	--
	Strongly Disagree	--	--

+ = Cycle II Directors.
* = Cycle I Directors of Validated Projects.

directors "strongly agree" that their project will be validated for Massachusetts Diffusion, forty-three percent (42.8%) agreed and eleven percent (10.7%) were neutral.

An Archive Review produced data relevant to the Massachusetts Validation Program success rate. These data are presented in Table 21.

Data indicated that more than half (52.5%) of the development projects funded under the Massachusetts Title IV-C program have been validated for diffusion. More projects were validated for diffusion during FY79 (20) than were validated during FY78 (12). A higher percentage of funded projects were validated for diffusion during FY79 (54.1%) than in FY78 (50.0%). In three regions the percentage of funded projects which were validated for diffusion exceeded the state average (52.5%). They were the Northeast Region (62.5%), Pittsfield/Springfield Regions (60.0%), and the Greater Boston Region (55.6%).

Data also indicated that thirty-two projects were validated for diffusion under the Title IV-C program in Massachusetts. More of these projects are located in the Greater Boston Region (10) and in the Pittsfield/Springfield Regions (9) than in any other regions.

Table 21

Massachusetts Validation Program FY78 and FY79

Region	CYCLE I (FY78)			CYCLE II (FY79)			TOTAL NUMBER VALIDATED PROGRAMS		
	No of Funded Projects	No. of Validated Projects	Percent	No. of Funded Projects	No. of Validated Projects	Percent	No. of Funded Projects	No. of Validated Projects	Percent
Greater Boston	6	2	33.3	12	8	75.0	18	10	55.6
Central Mass.	3	1	33.3	7	3	42.9	10	4	28.6
Northeast	3	1	33.3	5	4	80.0	8	5	62.5
Pittsfield/ Springfield	8	5	62.5	7	4	57.1	15	9	60.0
Southeast	4	3	75.0	6	1	16.7	10	4	40.0
Mass. Title IV-C Program (all regions)	24	12	50.0	37	20	54.1	61	32	52.5

The 1978 and 1979 Massachusetts Title IV-C Assessment Questionnaires examined the Cycle II directors' views of program success. In 1978, the directors were asked the question "To what degree of success, as defined by your own terms, do you anticipate your project will attain by the end of three years?" These data are presented in Table 22.

Twenty-seven percent (27.0%) of these directors anticipated that their projects will attain a "very high" degree of success after three years of project implementation; nearly sixty-eight percent (67.6%) indicated high, and five percent (5.4%) indicated medium.

In 1979, the question was revised and stated as "To what extent do you agree or disagree with the following statement: I believe that the project will be successful." These data are presented in Table 22.

Over sixty percent (61.8%) of the Cycle II directors indicated that they "strongly agree" that the project will be successful, thirty-five percent (35.3%) agreed, three percent (2.9%) had neutral feelings.

Table 22

Results of the Project Directors' Responses to Questionnaire Items Focused on Validation

Item	Type of Response	Cycle II Project Directors: Number and Percent of Responses Number	Percent
To what degree of success, as defined by your own terms, do you anticipate your project will attain by the end of three years?	1978		
	Very High	10	27.0
	High	25	67.6
	Medium	2	5.4
	Low	--	--
	Very Low	--	--
To what extent do you agree or disagree with the following statement "I believe that the project will be successful."	1979		
	Strongly Agree	21	61.8
	Agree	12	35.3
	Neutral	1	2.9
	Disagree	--	--
	Strongly Disagree	--	--

Evaluators

The Massachusetts Title IV-C Assessment Questionnaires included Categories of Investigation which examined the provision of Technical Assistance to project directors by project evaluators. These data are presented in Table 23.

The directors of Cycle II projects reported receiving varying amounts of technical assistance from their evaluators in both their first and second years of program implementation in the areas of proposal writing (38.9%, 1978), program planning (62.2%, 1978; 27.1% 1979), program implementation (51.4%, 1978; 35.2%, 1979), program evaluation (86.5%, 1978; 76.5%, 1979), training/skill development (43.2%, 1978; 35.3%, 1979), preparation for validation (69.4%, 1978; 58.8%, 1979), dealing with required procedures (64.9%, 1978; 38.2%, 1979).

The directors of Cycle I projects received assistance from their project evaluators in their first year of program implementation in proposal writing (33.3%), program planning (44.4%), program implementation (22.2%), program evaluation (66.7%), training/skill development (22.2%), preparation for validation (66.7%), and dealing with required procedures (55.6%).

The Cycle I directors of projects that were validated for diffusion reported receiving assistance in 1979 from their project evaluators with program

Table 23

The Project Directors' Response to the Question "How helpful has your program evaluator been with program planning, program implementation, program evaluation, training/skill development, preparation for validation, dealing with required procedures, preparing for diffusion, and identifying diffusion sites?"

	Percent of Responses Indicating "Helpful" or "Very Helpful" in each category					
	1978		1979			
Item	Cycle I	Cycle II	Cycle II	Cycle I Val. for Diffusion	Cycle I Val. for Dissem.	Cycle I Not Val.
Proposal Writing	33.3	38.9	--	--	--	--
Program Planning	44.4	62.2	27.1	43.4	58.4	0
Program Implementation	22.2	51.4	35.2	34.7	41.7	0
Program Evaluation	66.7	86.5	76.5	78.3	66.7	0
Training/Skill Development	22.2	43.2	35.3	30.4	41.7	0
Preparation for Validation	66.7	69.4	58.8	65.2	58.3	0
Dealing with Required Procedures	55.6	64.9	38.2	30.4	66.7	0

Table 23 (continued)

Item	Percent of Responses Indicating "Helpful" or "Very Helpful" in each category					
	1978		1979			
	Cycle I	Cycle II	Cycle II	Cycle I Val. for Diffusion	Cycle I Val. for Dissem.	Cycle I Not Val.
Preparing for Project Diffusion	--	--	--	21.7	--	--
Identifying Diffusion Sites	--	--	--	17.4	--	--

planning (43.4%), program implementation (34.7%), program evaluation (78.3%), training/skill development (30.4%), preparation for validation (65.2%), dealing with required procedures (30.4%), preparing for project diffusion (21.7%), and identifying diffusion sites (17.4%).

The directors of Cycle I projects that were only validated for dissemination reported receiving assistance from their project evaluators in the areas of program planning (58.4%), program implementation (41.7%), program evaluation (66.7%), training/skill development (41.7%), preparation for validation (58.3%), dealing with required procedures (66.7%).

The directors of projects that were not validated reported receiving no assistance from their evaluators.

The Massachusetts Title IV-C Assessment in 1979 examined the evaluation component of Cycle II projects. These data are presented in Table 24.

Fifty-one percent (50.9%) of the Cycle II project directors are reportedly evaluating student achievement, fifty percent (49.1%) are evaluating student attitude, twenty-five percent (25.5%) are evaluating other student changes (e.g., social, physical), thirty-six percent (36.4%) are evaluating student skill development, twenty-nine percent (29.1%) each, are evaluating personnel/teacher achievement and attitude, twenty

Table 24

The Project Directors' Response to the Question "What areas of your project are being evaluated?"

Item	Cycle II Project Directors: Number and Percent of Responses	
	Number	Percent
Student Achievement	28	50.9
Student Attitude	27	49.1
Other Student Changes (e.g., social, physical)	14	25.5
Student Skill Development	20	36.4
Personnel/Teacher Achievement	16	29.1
Personnel/Teacher Attitude	16	29.1
Personnel/Teacher Skill Development	11	20.0
Cost Effectiveness	11	20.0
Parent Attitude	13	23.6
Community Attitude	8	14.5

percent (20.0%) each, are evaluating personnel/teacher skill development and cost effectiveness, twenty-four percent (23.6%) are evaluating parent attitude, and fifteen percent (14.5%) are evaluating community attitude.

MDE Perceived Quality of Evaluation

Following is a description of the interview data concerning Massachusetts Department of Education perception of program evaluation. These data, gathered by an interviewer other than the investigator, identified the strengths of program evaluation in Massachusetts and included: State Education Agency (SEA) knowledge of projects; SEA knowledge of behavior of validation teams; director knowledge of SEA officials and the amount of technical assistance provided by them. The weaknesses of program evaluation were identified and include: inconsistency of evaluation message, attitudes across the state; lack of knowledge of formal evaluation skills; inaccessibility of state because of distance involved; loss of objectivity due to close identification of SEA officials to project; too little time to devote to this area. The SEA officials made the following recommendations: consistent technical assistance must be provided across the five Massachusetts Department of Education regions by the five State Education Agency officials;

the skill level of SEA staff, directors, evaluators, must be increased through training; statewide evaluation improvement; development of an evaluation/management system, creation of an evaluation support system.

Objective 2

Problems, based upon data obtained via surveys and document analysis, were identified, which contribute to the unsatisfactory evaluation outcomes of Title IV-C programs. Following is a listing of the identified problems:

1. The directors have an erroneously high perception of the quality of program evaluation (Table 8).

2. The directors of projects that were not validated for diffusion paid their program evaluator to assist them with project management. The directors of validated programs were more likely to utilize their program evaluators for program evaluation (Table 23).

3. The directors feel they know more about program evaluation than they do (Table 9).

4. Only half of the directors are reportedly measuring the achievement of the students involved with the project. Further, only a few evaluators are measuring teacher achievement or cost effectiveness (Table 24).

5. The evaluation designs submitted by the projects that were not validated for diffusion, did not yield data that demonstrated change (growth) in the client group (Table 12).

6. Evaluation designs were solely formative in nature (Table 15).

7. Evaluation designs did not produce conclusive evidence that the observed change was due to the project (Table 13).

8. Evaluation designs did not compare treatment vs. nontreatment conditions (Table 13).

9. The evaluators did not utilize a quasi-experimental design or, the evaluation designs were not sufficiently sensitive to provide such information (Table 13).

10. Evaluation designs did not include documentation that the treatment group is similar to control group (Table 13).

11. Planning for conducting the evaluation in adopter schools was not included (Table 13).

12. Data collection schedules were not included (Table 13).

13. Evaluation personnel and audiences were not included (Table 13).

14. Projects were under resourced to carry out adequate evaluation designs (page 112).

15. Project goals were too broad in scope (page 112).

16. Evaluation designs were focused on the Massachusetts Validation Process rather than on the goal of validation for national dissemination by the Joint Dissemination Review Panel (page 112).

17. Implementation of the evaluation design relied, in most instances, on unskilled persons (page 112).

18. Evaluation designs focused on small ends rather than on outcomes that will be educationally significant (page 113).

19. Evaluators utilized locally developed instruments that were not proven to be reliable or valid measures of the program (Table 15).

20. Test development procedures as well as test reliability and validity procedures were not included in evaluation designs (Tables 15 and 16).

21. Instrumentation utilized to measure specific objectives did not always "fit" (Table 16).

22. Few program evaluators included a clear description of the statistical procedures they intended to utilize in the evaluation. Further, those projects that were not validated for diffusion, rarely properly employed the statistics they used (Tables 18 and 19).

23. Program evaluators wrote objectives that addressed the key components of the project. However, the objectives were not complete or operational (Table 10).

24. Most projects that were validated for Massachusetts Diffusion are not being evaluated for presentation to the Joint Dissemination Review Panel (Table 20).

Objective 3

An evaluation plan which is tailored to resolve evaluation problems was identified. Guidelines were developed, which will convey to project directors and project evaluators the concept that evaluation designs for Title IV-C projects must be of such a quality that it would enhance a project's capability of being initially validated for diffusion by the Massachusetts Validation Process and subsequently validated for national

dissemination by the Joint Dissemination Review Panel. A complete description of the evaluation plan is presented in Chapter V.

Interpretation of the Results

The data for each of the study categories of investigation were previously discussed. The remainder of the chapter is dedicated to the interpretation of the results.

The Project Director

The directors of the Title IV-C projects which were included in the sample of this study are experienced educators who either helped to plan the project or wrote the proposal; were classroom teachers prior to being appointed as project director; have been in the school system between 1 and 5 years; have earned a Master's Degree.

Despite their sophistication, the directors have an erroneously high perception of the quality of their program evaluations. Further, the directors feel they know more about program evaluation than they actually do. Although not uncommon among practitioners who are on the cutting edge of change, these views make the directors particularly vulnerable to shoddy evaluation. To remedy this situation, a comprehensive evaluation training

program should be offered to the directors. The training program must allow the practitioner to see the different types of educational evaluation and measure them against state and federal guidelines for validation. With this knowledge, the directors can make clearer decisions about the use of evaluators and evaluation.

Program Objectives

A review of the data indicates that program goals were too broad in scope. Further, program evaluators wrote objectives that addressed the key components of the project. However, these objectives were neither complete or operational despite the fact that the directors reported in both 1978 and 1979 that they had revised their objectives.

Complete, specific and operational objectives are a necessary foundation to good program evaluation. In the first month of program implementation the directors should receive training in the area of program goal and objective writing giving the directors the control to write their own objectives or monitor an evaluator paid to write the objectives.

The Evaluation Design

Numerous weaknesses in the evaluation process were identified. First, many of the evaluation designs were solely formative in nature. This unfortunate circumstance provided directors and validators a description of what the project did, but not how well they did it. Lacking a summative evaluation it is impossible to judge the effectiveness of the project in the area of student change.

The evaluation designs submitted by the evaluators of projects that were not validated for diffusion, did not yield data that demonstrated change (growth) in the client group. In general, the evaluators did not produce conclusive evidence that the observed change was due to the project. Further, most evaluators did not compare treatment vs. non-treatment conditions or use any type of quasi-experimental evaluation design. When some type of comparison was included in the evaluation design, the evaluators did not include documentation that the treatment group was similar to the control group. The evaluators focused on small ends rather than on outcomes that would be educationally significant. Finally, evaluation designs did not include: plans for conducting the evaluation in future adopter schools; data collection schedules; evaluation personnel and audiences.

Data indicates that the evaluators either did not, or could not put forth an appropriate evaluation effort. The directors in their naivete contracted for inadequate evaluation. In addition to evaluation training, an evaluation management system, implemented by trained personnel, should be created and activated. This system would not only serve as an evaluation quality control system for state department of education administrators and auditors, but would also make individual technical assistance, above and beyond the training program available to the directors.

Instrumentation

Evidence of the inappropriate use of instrumentation in program evaluation was uncovered. Evaluators utilized locally developed instruments that were not proven to be reliable or valid measures of the program. Further, test development procedures as well as test reliability and validity procedures were not included in the evaluation designs. In addition, instrumentation chosen to measure specific objectives were not always accurate indicators of the objectives in question. Finally, sampling procedures were rarely correctly included in the evaluation design.

Evaluations lacking in instrument reliability and validity, produce no conclusive proof of project success or failure. Project directors must be made aware of the validation standards which require valid and reliable instrumentation to be used in any program evaluation. Training must provide the opportunity for directors to understand these concepts. The evaluation management system must retain a quality control over the hiring of a psychometrician to devise and validate instrumentation.

Statistics

Few program evaluators included in their designs a clear description of the statistical procedures they intended to utilize for the evaluation. Further, the evaluators of projects that were not validated for diffusion, rarely properly employed the statistical procedures they used. To correct for this, an evaluation management system needs to be put in place which does not allow a project director to hire an evaluator who has not proven his/her ability to correctly utilize statistics, and who cannot completely explain the proposed utilization of statistics in the evaluation design.

Validation

Most of the project directors believed that their programs would be validated for Massachusetts Validation. Approximately half were validated. Although this point of view is a healthy one for the change agent, in that it does not limit the innovation to failure, a training program which would make the validation standards clearer to project directors may help to remove false hopes early enough to correct for problems that limit the success of the program with validation. Understanding the validation process and its rewards may also help to alleviate the fact that the number of directors who felt that the Massachusetts Validation Process was an opportunity in 1978 decreased in 1979, while the number of directors who viewed the validation process as a hindrance dramatically increased in 1979.

To this date, no ESEA Title IV-C project in Massachusetts has been validated for federal dissemination. While thirty percent (30.0%) of the directors of projects that were validated for Massachusetts diffusion believe that their projects will be validated by the JDRP for federal dissemination, most Massachusetts validated projects are not being evaluated for presentation to the JDRP. Further, program evaluators focused on the Massachusetts Validation Process rather than on

the goal of validation for national dissemination by the JDRP. An evaluation management system, which assures that evaluators are designing evaluations sophisticated enough to meet the standards of the JDRP should be implemented.

Evaluators

Half of the program evaluators are measuring the achievement of the students involved with the project. Further, only a few evaluators are measuring teacher achievement or effectiveness. An evaluation design which does not measure student achievement, does not produce evidence of the most important nature. That evidence being, "Did this project produce a positive change in the children participating?"

The directors of projects that were not validated for diffusion paid their program evaluators to assist them with project management. The directors of validated programs were more likely to utilize their program evaluators for program evaluation. A training program designed to facilitate the directors' ability to be good consumers of evaluation should be offered early in the funding cycle.

A further review of the data indicates that: (1) the projects were under resourced to carry out

adequate evaluations; and (2) implementation of the evaluation relied in most instances on unskilled persons. In program evaluation this combination is deadly. These reasons more than any other account for the evaluation weakness previously indicated. Evaluation budgets that are under resourced allow only minimal effort on behalf of a trained evaluator, or disastrous results from an evaluator that is not trained. Evaluation budgets need to be adequate, directors need evaluation training in order to become good consumers of evaluation, and a quality control evaluation system must be implemented by Title IV-C personnel in order to insure high quality evaluations.

Massachusetts Department of Education

Massachusetts Department of Education (MDE) Title IV-C personnel indicated that the evaluation message from MDE personnel to directors and evaluators is inconsistent across the state. A lack of knowledge of formal evaluation skills was also identified. MDE personnel believe that evaluation training, technical assistance and an evaluation support system should be provided to the directors. Recommendations also included the creation of an evaluation management system.

Summary

Weaknesses in the various aspects of program evaluation were identified. The weaknesses in most cases, were directly related to the directors' knowledge of evaluation and ability to be good consumers of evaluation, inadequate evaluation budgets, untrained evaluators, and lack of management on behalf of Title IV-C personnel.

The data indicate that the directors' poor understanding of evaluation and/or an inadequate evaluation budget often result in the hiring of untrained evaluators; the result is usually a poor program evaluation. Changes in the ESEA Title IV-C program including (1) appropriate evaluation budgets evaluation training; (2) the implementation of an evaluation management system should result in quality program evaluations; and (3) management system is discussed in detail in Chapter V.

C H A P T E R V

THE EVALUATION IMPROVEMENT PLAN

Introduction

In Chapter IV several evaluation deficiencies were identified. These inadequacies as well as some suggested solutions were shared with the Title IV-C Coordinator and staff. At this time the Massachusetts Department of Education Title IV-C personnel pledged their support to the creation of an evaluation resource.

It was decided that resources must be identified or developed which will, by providing models and guidelines, structure project evaluation to assure that a conclusive demonstration of effectiveness is provided. Through assessment and consequent development of resources, the evaluations of Title IV-C projects will be extensively restructured.

The ultimate purpose is to assure that Title IV-C evaluations will conclusively determine student change and establish the cause of that change, as prescribed by Federal guidelines and as sought by the Joint Dissemination Review Panel.

The immediate purpose is to create an evaluation resource which will individualize both the evaluation efforts of Title IV-C projects and the training and assistance the projects receive.

In light of this, the investigator submitted a Title IV-C statewide proposal that went through a competitive screening process. The results were positive and the Evaluation Improvement Project was funded. An office was set-up as the Evaluation Resource Center.

The Comprehensive Evaluation Plan for Title IV-C in Massachusetts

In order to accomplish the task of improving the Title IV-C program evaluations in Massachusetts two stages were necessary. The first was a planning stage in which the evaluation goals and strategies of the Massachusetts Department of Education were operationalized. The second involved the development and implementation of a comprehensive evaluation improvement plan. A description of these two stages follows.

The Operationalization of MDE Evaluation Goals and Strategies

A conference was held in order to clarify the evaluation goals of Title IV-C in Massachusetts, and then to outline an evaluation improvement plan. It was

hoped that the results of this conference would give some basis for the Title IV-C Coordinator and the Evaluation Resource Center to plan, determine and institute desired procedures. These procedures would insure that the evaluation components of the Cycle IV projects would demonstrate conclusive evidence of effectiveness. The meeting participants included: Dr. Wayne Peters, Chairman, Title IV Advisory Council; Dr. John Reynolds, Coordinator, Massachusetts ESEA Title IV-C; Dr. Margaret Cassidy, Regional Program Officer, ESEA Title IV-C; Dr. Gordon Schimmel, Regional Program Officer, ESEA Title IV-C: Dr. William Allen, Director, Title IV-C Development Center; Dr. Ronald L. Nuttall, Laboratory for Statistical and Policy Research, Boston College; Nancy Taylor, U.S. Office of Education; Mary Seroski, M.E.C. Evaluation Project; Timothy Burns, M.E.C. Evaluation Project; and Dr. John Terry, Director, Project STYLE.

The results of the meeting determined that the goals of Title IV-C in Massachusetts include: (1) project validation for diffusion by the Massachusetts Validation Process, and (2) project validation for dissemination by the Joint Dissemination Review Panel. In order to enhance a Cycle III project's capability to reach the goals of Title IV-C, evaluation guidelines

that would span the life of the project should be adopted. A comprehensive evaluation plan for Title IV-C in Massachusetts was generated by the investigator and reported to the Title IV-C Coordinator and staff and to the Title IV Advisory Council.

The Comprehensive Evaluation Improvement Plan: A Description

The comprehensive Evaluation Improvement Plan for Title IV-C in Massachusetts was summarized in The Stages of Educational Evaluation in Innovative Projects and is presented in Table 25.

The Project Development Stages include: Stage I—Initiation; Stage II—Implementation; and Stage III—Outcomes. These stages were developed and researched from a project management point of view by the Rand Corporation and published in Federal Programs Supporting Educational Change (Rand, 1977).

In the Initiation Stage, evaluation should be oriented toward formative purposes. The goals of the evaluation should be to assist the project director in improving the program. The evaluation plan for objects in the initiation stage includes provisions for the Evaluation Resource Center to develop Formative evaluation kits. These should include questionnaires

Table 25

THE STAGES OF EDUCATIONAL EVALUATION IN INNOVATIVE PROJECTS

	YEAR I			YEAR II (Massachusetts Validation)			YEAR III (Massachusetts Validation)		
	JULY	AUGUST	SEPT.-JUNE	JULY	AUGUST	SEPT.-JUNE	JULY	AUGUST	SEPT.-JUNE
Phase I Initiation	Project Director Employed Development Phase Assessed	FORMATIVE EVALUATION IN PLACE	FORMATIVE EVALUATION	Development Phase Assessed	FORMATIVE EVALUATION Projects grouped in goal areas Psychometrician identified for necessary instrument development	Formative Evaluation SUMMATIVE INSTRUMENT DEVELOPMENT	Development Phase Assessed	Formative Evaluation SUMMATIVE EVALUATION IN PLACE	Formative Evaluation SUMMATIVE EVALUATION
Phase II Implementation	Project Director Employed Development Phase Assessed	FORMATIVE EVALUATION Projects grouped in goal areas Psychometrician identified for necessary instrument development	FORMATIVE EVALUATION SUMMATIVE INSTRUMENT DEVELOPMENT	Development Phase Assessed	Formative Evaluation SUMMATIVE EVALUATION IN PLACE	Formative Evaluation SUMMATIVE EVALUATION			Formative Evaluation SUMMATIVE EVALUATION
Phase III Outcomes	Project Director Employed Development Phase Assessed	FORMATIVE EVALUATION IN PLACE SUMMATIVE EVALUATION IN PLACE	Formative Evaluation SUMMATIVE EVALUATION						Formative Evaluation SUMMATIVE EVALUATION
SINGLE POPULATION	Project Director Employed Development Phase Assessed	Formative Evaluation in Place SUMMATIVE EVALUATION ASSISTANCE PROVIDED (pre-test/ design)	Formative Evaluation SUMMATIVE EVALUATION ASSISTANCE PROVIDED (pre-test/ design)	Development Phase Assessed	Formative Evaluation FINAL SUMMATIVE EVALUATION IN PLACE	Formative Evaluation SUMMATIVE EVALUATION			Formative Evaluation SUMMATIVE EVALUATION

Table 25 (continued)

Qualifications Needed in Order to be Placed in Each Evaluation Phase

Outcomes

- Measurable student learning objectives have been established.
- Project is known in system.
- Project director is knowledgeable of the project needs and of the decision-making dynamics of the community.
- Target population identified.
- System is supportive of project.
- Project Curriculum has been thoroughly developed and is currently in use.
- Project materials (products) are in existence.
- Project staff is experienced with project.
- Instrumentation for evaluation/validation are known and acceptable.

Implementation

- Project director is experienced with project.
- Project goals are NOT clear.
- Project Curriculum has been thoroughly developed and is currently in use.
- Target population identified.

OR

- Project director is NOT experienced with project.
- Project goals and/or objectives are clear.
- Project Curriculum has been thoroughly developed and is currently in use.
- Target population identified.
- any similar combination.

Initiation

- A project that is not yet in the IMPLEMENTATION Stage.

Table 25 (continued)

Single Population

- A project that will have a single target population over a three year period.
- Immediate Summative Evaluation assistance will be needed. The project should work with Title IV-C Coordinator, Title IV-C Regional Program Officers, and the Evaluation Resource Center in the selection of the project evaluator.

that the directors themselves could administer to various populations. This evaluation plan should be introduced by the Evaluation Resource Center's Evaluation Coordinator (the study investigator) and the Title IV-C Coordinator through training. It will be implemented by Regional Program Officers and Project Directors.

In the Implementation Stage, evaluation should be a combination of formative and summative evaluation. Formative evaluation will continue to provide information of use to the project dirctor in improving the program, while a strong summative evaluation is being designed. In training, provided by the Evaluation Resource Center, projects are grouped by content area (Phi Delta Kappa Goals). Standardized instruments are reviewed and selected. If no instrument is available, a psychometrician is contracted to develop instrumentation.

Evaluation should be summative in nature in the Outcomes Stage. Evaluation activities should be oriented toward gathering student effectiveness data and achieving statewide and national validation. At this point, formative evaluation is voluntary. Summative Evaluation Kits should be developed by the Evaluation Resource Center. Summative evaluation should be implemented immediately.

Finally, the plan included provisions for the Coordinator of Evaluation to develop a scale which would measure a project's stages of development against criteria indicative of a mature project, as well as a PDK goal area check list.

The Project Scale

A Project Scale which would measure the developmental stage of a project was developed in June and July of 1979 by the study investigator. Attached to the Project Scale was a Project Scale Outline, which listed project management indicators for each of the three developmental phases for clarification purposes. The Project Scale and the Project Scale Outline were presented to the Title IV-C Coordinator and the Title IV-C staff for approval in July 1979. Following the meeting, revisions were made on both documents and Working Copy II of the Project Scale and the Project Scale Outline were developed and presented to the Title IV-C Coordinator and staff and approved in late July 1979. On 20 July 1979 the Project Scale was field tested in the Central Massachusetts Region. At this time, the Title IV-C Coordinator, the Central Massachusetts Regional Program Officer, and the Coordinator of Evaluation, of the Evaluation Resource Center met with each Cycle III project director and superintendent,

as well as project staff. A package consisting of The Stages of Educational Evaluation in Innovative Projects and Working Copy II of the Project Scale and Project Scale Outline were presented and administered. As a result of the field test, minor changes in the physical appearance of the Project Scale and the Project Scale Outline were made and the Instruments were finalized. In August and September of 1979 the Evaluation Resource Center presented the Comprehensive Evaluation Plan to Cycle III Project Directors in the Boston I, Boston II, and Northeast regions. Once again the Title IV-C Coordinator, the Regional Program Officer, and the Coordinator of Evaluation Services of the Evaluation Resource Center, met with Cycle III project directors and superintendents. Packages consisting of The Stages of Educational Evaluation in Innovative Projects, the Project Scale and the Project Scale Outline were presented and administered. In the Southeast and Springfield/Pittsfield regions, the Regional Program Officer administered the Project Scale. A completed Project Scale for each Cycle III project is on file at the Evaluation Resource Center. The implementation of the Project Scale served not only as an indicator of the developmental level of a project for the director, but also served as a needs assessment for the Evaluation Resource Center. Thirty-five of the thirty-seven Cycle

III projects are in the Initiation Stage. The Coordinator of Evaluation Services of the Evaluation Resource Center planned the evaluation training for the director of a project in the Initiation Stage. The Project Scale and the Project Scale Outline are presented in Appendix D.

The Formative Evaluation Manual

As necessitated by the Comprehensive Evaluation Plan for Title IV-C in Massachusetts, the Evaluation Resource Center staff took on the huge task of developing the Formative Evaluation Manual. In the summer of 1979 Evaluation Resource Center staff along with the Title IV-C Coordinator conceptualized the contents of the Manual in the form of an outline. In August 1979, Michael S. Walker and Thomas E. Wolf wrote the first draft of A Formative Evaluation Manual. Evaluation Resource Center staff and the Title IV-C Coordinator examined the document and suggested revisions. Wolf and Walker produced an updated document which was approved by the Evaluation Resource Center staff and the Title IV-C Coordinator. In August 1979 the Evaluation Resource Center hosted a meeting in order to present the manual design and concept to the Title IV-C staff. Participants included: Dr. John Reynolds, Title IV-C Coordinator; Dr. Gordon Schimmel, Southeast Regional Program Officer; Dr. Michael Mayo,

Greater Boston Regional Program Officer; Dr. Margaret Cassidy, Greater Boston Regional Program Officer; Charles Radlo, Central Massachusetts Regional Program Officer; Maria Grasso, Northeast Regional Program Officer; Thomas E. Wolf and Michael S. Walker, Authors of A Formative Evaluation Manual; Tim Burns and Mary Seroski of the Evaluation Resource Center and Maurice Smith, M.E.C. Liaison. The authors made a presentation on the manual design. Approval was received and the manual went to print.

An exposure draft of A Formative Evaluation Manual was field tested with Cycle III project directors from September through December 1979. As a result of the field test, changes in the manual were requested by the Title IV-C Coordinator and Evaluation Resource Center staff. A Formative Evaluation Manual has been updated by Wolf and Walker and is now ready for printing.

Evaluation Training

In order to implement the Comprehensive Evaluation Plan for Title IV-C in Massachusetts, an Evaluation training program was implemented. The Coordinator of Evaluation Services planned an evaluation training program which was made an integral part of monthly project director meetings.

Most of the Cycle III projects were in the Implementation Stage, thus training in Formative Evaluation was planned. On 7 September 1979 the Coordinator of Evaluation Services sent a memo to all Cycle III directors with a Draft Exposure of A Formative Evaluation Manual. The memo explained the evaluation training, and the manual. On 12 September 1979 authors Wolf and Walker sent a memo to all Cycle III project directors giving them an assignment for the Formative Evaluation training session that would take place at the September Project Director meeting.

The Formative Evaluation training began at the Cycle III project director meeting on 19 September 1979. Thomas Wolf and Michael Walker presented the first of three workshops on Formative Evaluation. On the same day, the Coordinator of Evaluation Services made a brief introductory presentation on Summative evaluation for evaluation. At this time two documents from the library at the Evaluation Resource Center were lent to each of the Cycle III project directors. The documents are, Massachusetts Validation Process, Overview and IDEABOOK, The Joint Dissemination Review Panel.

Evaluation training was continued at the 23 October 1979 project director meeting. At this point, in accordance with the Title IV-C Evaluation Plan, the

directors were well on their way to planning their own Formative Evaluation. To advance the directors progress the second of three workshops on Formative Evaluation was presented by Wolf and Walker. In order to prepare the directors to move into the Initiation phase, Dr. John Reynolds, Title IV-C Coordinator, made a presentation on summative evaluation. The Summative Evaluation Plan as presented to the directors and their superintendents in the summer of 1979 was reiterated. Dr. Reynolds explained the Summative Evaluation of Title IV-C projects in Massachusetts and the Summative Evaluation Management System.

On 28 November 1979 Wolf and Walker presented the final workshop on Formative Evaluation. The project directors now, are capable of planning and implementing a formative evaluation.

The next step in the evaluation process is the identification of instrumentation. To assist the directors with this awesome chore, the Coordinator of Evaluation Services distributed bibliographies of instruments that are part of the library housed in the Evaluation Resource Center.

At the 23 January 1980 project director meeting, the Coordinator of Evaluation Services presented an evaluation progress report to the directors. Shortly after the new year, many of the directors were preparing

to hire a psychometrician. Dr. Ronald Nuttall, Professor of Educational Research, Measurement, and Evaluation, Boston College, presented at the 26 March 1980 project director meeting. Dr. Nuttall discussed the utilization of a psychometrician.

At the 13 May 1980 project director meeting, the Coordinator of Evaluation Services was available to each director for personal evaluation consultation.

At the conclusion of the second year of program implementation, these projects will be eligible to apply for Massachusetts Validation for Diffusion. The training for year two was planned to equip the directors with the skills necessary to compete in the validation process. At the 16 September 1980 project director meeting, the Title IV-C Coordinator presented the Summative Evaluative Management System, once again. In addition, Dr. Ronald Nuttall, Professor of Educational Research, Measurement and Evaluation, Boston College, presented the evaluation standards for validation.

At the 21 October 1980 meeting, the Massachusetts Department of Education Title IV-C staff presented the states' procedures and requirements for validation.

The Coordinator of Evaluation Services, of the Evaluation Resource Center, was available at each meeting to provide individual technical assistance.

Technical Assistance

Technical Assistance is provided to project directors and the Title IV-C staff by the Coordinator of Evaluation Services on an ongoing and individual basis. This has been a highly successful practice.

The Summative Evaluation Management System

The Massachusetts Validation Process Workshop

On 25 September 1979 the Massachusetts Validation Process Workshop was hosted by the Evaluation Resource Center. The featured presenter at the workshop was Dr. William C. Wolf, Jr., Professor, University of Massachusetts. In suggesting possible revisions of the Massachusetts Validation Process, Dr. Wolf proposed that a Summative Evaluation Management System would give direction to projects seeking validation. Dr. John F. Reynolds, Title IV-C Coordinator and the study investigator developed an E.S.E.A. Title IV-C Summative Evaluation Management System and check-list. This system is being implemented by the Coordinator of Evaluation Services. The system makes it possible for a director to easily plan a comprehensive evaluation system by following 12 steps. The steps include:

(1) Student goals defined, (2) major objectives in terms of student results defined, (3) existing standardized tests reviewed, (4) psychometrician sought, (5) psychometrician selected and contracted, (6) instruments acceptable, (7) summative evaluator sought, (8) summative evaluator selected and contracted, (9) interim evaluation progress reports submitted, (10) data revised on validation request, (11) validation request submitted, and (12) Validation Team Report received. The Summative Evaluation Management System is presented in Appendix D.

This Summative Evaluation Management System was presented to the directors at the 23 October 1979 project director meeting. A record of the progress of each Cycle III project, with the Summative Evaluation Management System, is kept on file at the Evaluation Resource Center.

The system delivers the control of program evaluation to the Massachusetts Department of Education in that, an audit exception may be taken to any evaluation contract that the Title IV-C Coordinator does not approve.

The Summative Evaluation Resource

The purpose of the resource is to assist project evaluators and project directors by providing model approaches to the various stages of the evaluation task. The resources include sections dealing with instrumentation, research models, statistics and submission formats. Although the Summative Evaluation Resource is not scheduled to be completed until the Spring of 1982, a three year schedule has been compiled and is progressing according to the schedule. Following is an explanation of progress to date.

1. Instrument File

    a. All instrumentation appropriate for use by Title IV-C evaluators has been identified.

    b. Approximately 400 tests have been included in the library at the Evaluation Resource Center, located in Amherst, Massachusetts.

    c. The Coordinator of Evaluation at the Evaluation Resource Center and the University of Massachusetts Library staff have designed a cooperative system that make the tests in the library at the University of Massachusetts available to Title IV-C project directors.

    d. A professional librarian has catalogued the tests.

    e. The card catalogue has been typed and compiled.

    f. A check-out system is in place.

2. Research Models and Statistics

   a. The library houses approximately 140 books including two (2) books that deal exclusively with Research Models and twenty (20) additional books that deal with measurement. In addition, twelve (12) books on statistics are included in the library.

   b. A professional librarian has catalogued the texts.

   c. The card catalogue has been typed and compiled.

   d. A check-out system is in place.

3. File of Massachusetts Validation and JDRP Submissions

   a. Cycle I and Cycle II Massachusetts Validation Submissions have been compiled at the Evaluation Resource Center.

   b. Several JDRP Submissions have been placed on file at the Evaluation Resource Center. Sy Rubeck, Secretary of the Joint Dissemination Review Panel, has pledged his support in expanding this file.

Orientation Conferences for Project Directors

Conferences for project directors on the use of evaluation have been held at the September, October, November, and January project director meetings. In addition, privately scheduled conferences with project directors and the Coordinator of Evaluation have occurred.

Summary

The Comprehensive Evaluation Plan for Title IV-C in Massachusetts was developed, and includes three stages. The Project Development Stages include: Stage I—Initiation, in which evaluation should be oriented toward formative purposes; Stage II—Implementation, in which evaluation should be a combination of formative and summative evaluation; and Stage III—Outcomes, in which evaluation should be summative in nature.

The plan included provisions for the development and utilization of a project scale which would measure the development stage of a project. The scale was utilized in August of 1979 with the directors of Cycle III projects and will be utilized in September of 1980 with the directors of Cycle IV projects. Use of the project scale allowed the directors to gain insight about the maturity level of their projects as well as afford the Massachusetts Department of Education the opportunity to structure the development of the evaluation component of the program.

In order to implement the Comprehensive Evaluation Plan for Title IV-C in Massachusetts, as well as compensate for the directors' lack of knowledge in program evaluation, an evaluation training program was implemented. Training in formative evaluation, validation,

instrumentation, and summative evaluation was provided during FY79. Further, training for summative evaluation for validation is planned for FY80. In addition technical assistance was provided to the directors by the investigator on an ongoing and individual basis.

A summative evaluation resource that includes an instrument file, evaluation library, and a file of Massachusetts Validation and JDRP submissions has been compiled and is available for use by directors and evaluators. The resource will be enlarged in FY80. Further, a file of Summative Evaluators and psychometricians is currently being developed. A summative evaluation manual is in the planning stage.

CHAPTER VI

SUMMARY, CONCLUSIONS, AND IMPLICATIONS FOR FURTHER RESEARCH

Introduction

The major objectives of the study were: (1) to systematically study the state of Massachusetts' Title IV-C operation in order to document evaluation policies and practices; (2) to identify problems, based upon data obtained via surveys and document analysis, which contribute to the program's unsatisfactory evaluation outcomes; and (3) to offer an evaluation plan which is tailored to resolve evaluation problems identified within the state's Title IV-C program.

In Massachusetts, over 100 school systems have been funded to find creative solutions to local problems under Title IV Part C of the Elementary and Secondary Education Act of 1965. The funds have enabled local education agencies to enrich the school curriculum in many of the priority areas of the Massachusetts Board of Education. It is the intent of the legislation to provide "seed money" to a school system to test a new solution to a local problem, in order to identify

successful ventures through validation by the Joint Dissemination Review Panel (JDRP), resulting finally in the national dissemination of the innovation. Unhappily, not one of the Massachusetts Title IV-C projects have been presented to the JDRP. What has sadly unfolded is that these federally-funded projects cannot measure-up to established validation procedures. Millions of dollars have been invested in innovative education projects in Massachusetts. The return on this investment in the form of validated projects is not impressive.

A deep concern for this unfortunate state of affairs motivated the conceptualization of a descriptive study of Massachusetts Title IV-C program efforts. Given a set of Title IV-C projects funded in Massachusetts, efforts were made to ascertain why E.S.E.A. Title IV-C projects, which are funded and expedited within the state of Massachusetts, routinely fail to measure up to validation guidelines established by the state and federal governments, and, to offer remedies, based upon data obtained, that are likely to resolve identified deficiencies.

This chapter will discuss each of the study purposes in relationship to methodology, results and conclusions. This is followed by implications for further research.

Summary and Conclusions

Objective 1

The first component of the study involved the systematic investigation of the Massachusetts Title IV-C operation in order to document evaluation policies and practices. This was accomplished by designing a comprehensive study permitting data to be gathered from every available source.

Six data collection procedures including the Massachusetts Title IV-C Assessment, Massachusetts Validation Process, Archive Review, Interviews, Evaluation Design Review, and Assessment of Massachusetts Evaluation Needs, were utilized in order to study the state of Title IV-C in Massachusetts. Examination of the many categories of investigation enabled the investigator to document policies and practices and to identify problems which contribute to non-validation.

Massachusetts Title IV-C Assessment

The Title IV-C program in Massachusetts is assessed state-wide in the spring of each year. The assessment is designed to monitor the progress of Title IV-C in Massachusetts, and to satisfy the specific data needs of the Massachusetts Title IV-C Coordinator and staff. Massachusetts Title IV Advisory Council, and the United States Office of Education.

A series of questionnaires were employed in order to gather these data. The questionnaires were developed by the investigator according to the standards for questionnaire construction set forth by Kornhauser and Sheatsley (Selitz, 1976). The questionnaire development process included: needs assessment, conceptualization of assessment design, outline, ordering and analysis of topics for content and form; question development, question analysis, pretest, and revision.

The questionnaires were administered to the directors of projects that received initial funding in 1976 (Cycle I projects) and 1977 (Cycle II projects). Cycle I project directors completed a questionnaire in the spring of their second year of project implementation (1978) and in the spring of the third year of project implementation (1979). Cycle II project directors completed a questionnaire in the spring of their first year of project implementation (1978) and in the spring of the second year of project implementation (1979). In 1978, twenty-four (24) Cycle I and thirty-six (36) Cycle II directors received and returned questionnaires. In 1979, twenty-four (24) Cycle I and thirty-six (36) Cycle II directors received questionnaires. Questionnaires were returned by twenty-one (21) Cycle I and thirty-five (35) Cycle II directors.

The data were computer analyzed utilizing the Statistical Package for the Social Sciences (SPSS). Descriptive statistics were obtained and presented as director perception.

The Massachusetts Title IV-C Assessment Questionnaire provided data relative to the director's perception of project evaluation efforts. The Categories of Investigation include, Project Director, Program Objectives, Validation, and Evaluators.

Massachusetts Validation Process

The Massachusetts Validation Process is designed to validate worthy innovations. The process is accomplished by an on-site visit by a validation team. The Cycle I project directors had the opprotunty to apply for validation in the spring of 1978. The Cycle II projects were eligible for validation in the spring of 1979.

Following an on-site validation visit to a Title IV-C project, each validator completed a questionnaire. The questionnaire gathered data relevant to program evaluation. The questionnaire was developed by the investigator according to the standards for questionnaire construction set by Kornhauser and Sheatsley (Selitz, 1976). The development procedures included the

conceptualization of an assessment design followed by the delineation of topics in an outline. Once in outline form the topics were: ordered, examined for the most appropriate psychological sequence from the standpoint of the respondent, analyzed for content and form, and put into question form. The number of questions necessary on each item was considered. The questionnaires were then pretested, revised and administered.

One hundred thirty-five (135) questionnaires were distributed to the validators serving the nineteen (19) Cycle I projects and twenty-six (26) Cycle II projects who applied for Massachusetts Validation. The validators returned ninety-nine (99) questionnaires.

The Massachusetts Validation Process Questionnaire produced data relevant to program evaluation. The Categories of Investigation included evaluation design, statistics, and instrumentation.

Archive Review

An Archive Review was conducted. The document reviewed were the Massachusetts Validation Team Reports.

The Massachusetts Validation Team Reports are documents completed by the Validation Team members following each onsite visit. The reports include the Team's judgments about the effectiveness of the project as well as comments and suggestions.

Data relevant to the rate of project success with the Massachusetts Validation Process was provided.

Interviews

Interviews were conducted with a sample of the Massachusetts Department of Education Title IV-C Administrators. These interviews were conducted by an evaluator other than the study investigator. The interviews emphasized the acquisition of data about opinions and views of Massachusetts Title IV-C personnel on the perceived quality of program evaluation.

An interview questionnaire was developed according to the standards for questionnaire construction set by Kornhauser and Sheatsley (Selitz, 1976). Questionnaire construction included defining the topics in the form of an outline. Once in outline form, the topics were ordered, analyzed for content and form and organized in the form of questions. The questionnaire was pretested and revised.

The interviews were conducted by an evaluator other than the study investigator. The use of an external evaluator assured that the opinions of this investigator were not being projected on to those being interveiwed. The interviews were conducted in accordance with the standards for interviewing set by Kornhauser and Sheatsley

(Selitz, 1976). An interview questionnaire was completed by the interviewer following each interview.

The sample included the Massachusetts Title IV-C Coordinator and five (5) of the six (6) members of the Title IV-C staff (N=6). The sixth staff member was out of the country during the interview period.

The interviews focused upon the opinions and views of Massachusetts Title IV-C personnel on one Category of Investigation, the perceived quality of program evaluation.

Evaluation Design Review

Independent evaluators who have had experience with the JDRP and who are considered to be experts in educational evaluation reviewed Title IV-C program evaluation designs for worthiness. The evaluators who participated in the review were Louis Aikman, Ph.D., Boston University; Ronald Nuttal, Ph.D., Boston College; David J. Rosen, Ed.D., University of Massachusetts; Gene Mulcahy, Ed.D., Hartford Public Schools; Robert Algozzine, University of Florida; Martin J. Higgins, Ph.D., West Chester State College.

The sample included thirty-two (32) Title IV-C projects that received initial funding in 1977. The Evaluation Design Review was made available to the thirty-seven (37) directors of Title IV-C projects then,

in their first year of project implementation. The sample includes all the directors who wanted the service.

The evaluation design review provided data relevant to six Categories of Investigation, including program objectives, evaluation design, statistics and instrumentation.

Assessment of Massachusetts Evaluation Needs

A conference was held to identify the needs of Title IV-C projects in Massachusetts relative to program evaluation for validation. Participants included United States Office of Education staff, Massachusetts Department of Education staff, National Diffusion Network staff, validators, project directors, evaluation experts, and the study investigator.

These data reflect the degree to which Massachusetts Title IV-C projects have attained the evaluation standards prescribed by the Joint Dissemination Review Panel and the Massachusetts Validation Process. Information relevant to one category of investigation, evaluation design was produced.

Objective 2

The six data collection procedures enabled the investigator to study the operation of the Title IV-C program in Massachusetts. In order to carry out the second study objective, problems which contribute to the program's unsatisfactory evaluation outcomes were identified. This information was grouped to point out specific weaknesses in formative evaluation and summative evaluation, with an emphasis on instrumentation, research design, statistics, and success at validation. Following is a summary of the study results and conclusions based upon the specific evaluation inadequacies which were identified.

Project Director

Data indicated that although the directors of Title IV-C projects were sophisticated administrators, they felt they knew more about program evaluation than they actually did. This unfortunate situation makes the directors particularly vulnerable to shoddy evaluation.

Project director training in program evaluation should be provided in order to acquaint the directors with the theory of program evaluation. This training should result in the acquisition of the skills necessary to adequately manage and design a sufficient program evaluation for validation.

Program Objectives

Complete, specific, and operational objectives are a necessary foundation for good program evaluation. However, a review of the data indicates that program goals were too broad to scope and that objectives were neither complete or operational.

The project director is the heart of an educational innovation. It is this individual that most closely understands what the final impact of the program should be. In this light, it must be the responsibility of the project director, not the hired evaluator, to write the project objectives as adequate reflections of the innovation itself. In order to insure that the objectives are complete, specific, and operational, training should be provided to the directors which would enable them to carry out this most basic step in the evaluation process.

Evaluation Design

Numerous weaknesses in the evaluation process were identified. First, many of the evaluation designs were solely formative in nature. The evaluation designs submitted by the evaluators of projects that were not validated for diffusion, did not yield data that demonstrated change (growth) in the client group. In general, the evaluators did not produce conclusive evidence that

the observed change was due to the project. Further, most evaluators did not compare treatment vs. non-treatment conditions or use any type of quasi-experimental evaluation design. When some sort of comparison was included in the evaluation design, the evaluators did not include documentation that the treatment group was similar to the control group. The evaluators seemed to focus upon inconsequential ends rather than on outcomes that would be educationally significant. Finally, evaluation designs did not include: plans for conducting the evaluation in the future in adopter schools; data collection schedules; or evaluation personnel and audiences.

Conferences should be held for project directors and their evaluators in order to acquaint them with the United States Office of Education validation standards for evaluation design. Further, the Massachusetts Department of Education should implement an evaluation management system which would monitor the evaluation efforts of each Title IV-C project in order to insure the appropriate use of evaluation design.

Instrumentation

Evidence of the inappropriate use of instrumentation in program evaluation was uncovered. Evaluators utilized locally developed instruments that were not proven to be reliable or valid measures of the program.

Further, test development procedures as well as test reliability and validity procedures were not included in the evaluation designs. Instrumentation chosen to measure specific objectives were not always accurate indicators of the objective in question. Finally, sampling procedures were rarely correctly included in the evaluation designs.

An evaluation library which would include resources relevant to the appropriate uses of instrumentation for program evaluation should be compiled. An instrument file which would include all published instrumentation appropriate for use in Title IV-C evaluations should be made easily accessible to project directors and evaluators.

Statistics

Few program evaluators included in their designs a clear description of the statistical procedures they intended to utilize for the evaluation. Further, the proper usage of statistical procedures was rare in the case of projects that were not validated for diffusion.

An evaluation management system should be put in place which requires prospective evaluators to submit resumes, samples of previous work, and other information in support of their ability to perform each of the technical aspects of evaluation. The Massachusetts Department of Education Title IV-C staff should implement

a policy which would prohibit the employment of an evaluator who does not have the capabilities to determine the ultimate effectiveness of a program.

Validation

To this date no ESEA Title IV-C project in Massachusetts has been validated for federal dissemination. While thirty percent (30%) of the directors of projects that were validated for Massachusetts diffusion believed that their projects would be validated by the Joint Dissemination Review Panel (JDRP), most Massachusetts validated projects are not being evaluated for presentation to the JDRP. Further, program evaluators focused on the Massachusetts Validation Process rather than on the goal of validation for national dissemination by the JDRP.

Through training and technical assistance, project directors and evaluators should clearly understand the validation policies of the United States Office of Education as administered by the Joint Dissemination Review Panel.

Evaluators

A further review of the data indicates that: (1) projects were under resourced to carry out adequate evaluations; and (2) implementation of the evaluation relied in most instances on unskilled persons.

Current Title IV-C policy indicates that a director may spend ten percent (10.0%) of the total project budget for the purpose of evaluation. This amount may be sufficient for a project funded in the range of $90,000.00, but is certainly not adequate for a project funded in the range of $13,000.00. To accommodate for this, project directors must receive training in the various methods of acquiring the local support necessary to ensure needed financial support. Further, individual technical assistance should be available to project directors at each level of program implementation in order to insure the appropriate use of evaluation.

Objective 3

This portion of Chapter VI presents a description of the evaluation plan which was designed to resolve the evaluation problems identified within the state's Title IV-C program.

The study results and conclusions were shared with the Title IV-C Coordinator and staff. At this time the Massachusetts Department of Education Title IV-C personnel pledged their support to the creation of an evaluation resource.

It was decided that resources must be identified or developed which will, by providing models and guidelines, structure project evaluation to assure that a conclusive demonstration of effectiveness is provided. Through assessment and consequent development of resources, the evaluations of Title IV-C projects are being extensively restructured.

The ultimate purpose is to assure that Title IV-C evaluations will conclusively determine student change and establish the cause of that change, as prescribed by Federal guidelines and as sought by the Joint Dissemination Review Panel.

The immediate purpose is to create an evaluation resource which will individualize both the evaluation efforts of Title IV-C projects and the training and assistance the projects receive.

As a result, the investigator submitted a Title IV-C statewide proposal which was funded as the Title IV-C Evaluation Resource Center.

A comprehensive evaluation improvement plan was developed and implemented with the Cycle III directors. The evaluation plan is developmently oriented and includes: Stage I—Initiation, Stage II—Implementation, and Stage III—Outcomes.

In the Initiation Stage, evaluation should be oriented toward formative purposes. The goals of the evaluation should be to assist the project director in improving the project.

The Implementation State should be a combination of formative and summative evaluation. Formative evaluation will continue to provide information of use to the project director in improving the program while a strong summative evaluation is being designed.

Evaluation should be summative in nature in the outcomes stage. Evaluation activities should be oriented toward gathering student effectiveness data and achieving Massachusetts and Federal validation.

The plan includes provisions for the administration of the project scale, which indicates the developmental stage of a project. Evaluation training as well as individual technical assistance is being provided to the directors throughout the funding period.

A summative evaluation management system was developed and is being implemented. The system allows a director to easily plan a comprehensive evaluation system, while allowing the Massachusetts Department of Education Title IV-C personnel to retain control of program evaluations.

An evaluation library containing over 400 tests, 140 books, and a file of Massachusetts Validation and JDRP submissions has been made available to the directors.

The results of the study are optimistic. The Evaluation Improvement Project was funded for three years to develop and implement a Summative Evaluation Management System for Title IV-C projects in Massachusetts. The directors of Cycle III projects, who recieved initial funding in 1979, received the additional evaluation training and technical assistance. Those projects that will receive initial funding in 1980 will also receive the evaluation training and technical assistance. The ultimate criteria of success will occur in the future when the number of JDRP validated projects from Cycle III is known. Optimism must be noted in the fact that even though these Cycle III projects will not face Massachusetts Validation until the summer of 1981, the evaluation designs that have been reviewed by the investigator are sufficiently sophisticated to make an adequate attempt for submission to the Joint Dissemination Review Panel. One additional measure of the success of the Evaluation Improvement Project, is the fact that the Massachusetts Validation Process standards were thoroughly revised and strengthened, due to the high quality of evaluations that are being produced.

One unhappy result has been noted. In June of 1979, several directors criticized the summative evaluation management system for taking the control of one aspect of program management away from them. This mood of the directors must be avoided by all means. The great risk that these unhappy directors may "give-up" trying to evaluate their programs for validation must be avoided. Adjustments must be made in the evaluation management system to account for this.

The early success of this program is in part due to the Title IV-C Coordinator and staff whose understanding allowed the creation and implementation of the summative evaluation management system. The fact that the Title IV-C Coordinator and staff took ownership of the concept studied in this dissertation was a decisive factor in success to this point. If the state had not backed the Evaluation Improvement Project, the summative evaluation management system may very well have gone along unnoticed and unimplemented.

Implications for Further Research

Following are recommendations for further research:

1. Changes in evaluation policy implicated by the model put forth in this dissertation should be studied in more detail. This study concluded that measurement techniques were in the past, inappropriately used. The measures

that are currently in place to correct for this should be examined. The quality of existing as well as newly developed instrumentation must be assessed. This assessment must not be an exclusive comparison of prior usage in Massachusetts as any change would indicate a quantum jump. Current use of instrumentation must be compared against IVD and JDRP guidelines as well as standards set by other states.

2. A review of the summative evaluation designs of Cycle III projects should be conducted in the spring of 1980. The evaluation design review should be similar in form to that which was offered to Cycle II projects. The results should be compared.

3. The Cycle III success rate with the Joint Dissemination Review Panel should be studied.

4. The evaluation management system has put new demands on the budgets of Title IV-C projects. Even though these demands have increased, the percentage of program funds available for evaluation have not. The feasibility of implementating a comprehensive evaluation with the limited program funds available for evaluation (5% of total project budget) should be examined.

5. The appropriate inclusion of criterion-referenced tests or norm-referenced tests in the evaluation process, as well as the budget limitations on test development must be investigated.

6. A cross-state study designed to examine the evaluation procedures of Departments of Education in other states which have the validation success rate of Florida and New Jersey is needed.

7. Further study involving the implementation of the Evaluation Improvement Project in another state is necessary to ascertain if the project is sufficiently global to make validation and diffusion possible.

8. A follow-up study is needed to determine if the summative evaluation management system had any lasting effects on both the projects involved and the policy of the Massachusetts Department of Education.

B I B L I O G R A P H Y

Allen, M. J., & Yen, W. M. Introduction to Measurement Theory. Montery: Brooks/Cole Publishing Company, 1979.

Alkin, M. C. Evaluation Theory Development. In R. Worthen & W. Sanders (Eds.), Evaluation: Theory and Practice. Belmont: Wadsworth, Inc., 1973.

American Psychological Association. Standard for Educational and Psychological Tests. Washington, DC: Author, 1974.

Anastasi, A. Psychological Testing. New York: Macmillan Inc., 1976.

Barnhard, C. L. (Ed.). The American College Dictionary. New York: Random House, 1969.

Bernstein, I. N. Validity Issues in Evaluation Research. Beverly Hills: SAGE, 1976.

Bloom, B. S., Hastings, J. T., & Madaus, G. F. Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill, 1971.

Bloom, B. S. Taxonomy of Eduational Objectives, Handbook I: Cognitive Domain. New York: Longman, 1956.

California State Department of Education. Program Evaluation Guide. Princeton: Educational Testing Service, 1977.

Campbell, D. T., & Stanley, J. C. Experimental and Quasi-Experimental Designs for Research. Chicago: Rand McNally, 1963.

Campbell, D. T., & Erlebacher, A. How Regression Artifacts in Quasi-Experimental Evaluations can Mistakenly Make Compensatory Education Look Harmful. In M. Guttentag & E. L. Struening (Eds.), Handbook of Evaluation Research (Vol. 1). Beverly Hills: SAGE, 1975.

Cohen, J. Multiple Regression as a General Data—Analytic System. In M. Guttentag & E. L. Struening (Eds.), Handbook of Evaluation Research (Vol. 1). Beverly Hills: SAGE, 1975.

Cronbach, L. J. Course Improvement Through Evaluation. Teachers College Record, 1963, 64, 672-783.

Cronbach, L. J. Course Improvement Through Evaluation. In R. Wothen & G. Sanders (Eds.), Educational Evaluation: Theory and Practice. Belmont: Wadsworth, 1973.

Davis, H. R., & Salasin, S. E. The Utilization of Evaluation. In M. Guttentag & E. L. Struening (Eds.), Handbook of Evaluation Research (Vol. 1). Beverly Hills: SAGE, 1975.

Deming, W. E. The Logic of Evaluation. In M. Guttentag & E. L. Struening (Eds.), Handbook of Evaluation Research (Vol. 1). Beverly Hills: SAGE, 1975.

DuBois, P. H. A History of Psychological Testing. Boston: Allyn and Bacon, 1970.

Eber, H. W. Multivariate Methodologies for Evaluation Research. In M. Guttentag & E. L. Struening (Eds.), Handbook of Evaluation Research (Vol. 1). Beverly Hills: SAGE, 1975.

Eisner, E. W. The Perceptive Eye: Toward the Reformation of Educational Evaluation. Paper presented at the annual meeting of the American Educational Research Association, Washington, DC, 1975.

Ford, J. J., & Williams, M. K. Educational Manager's Guide to Project Planning: A Systematic Approach. Andover: The NETWORK, 1978.

Foster, R. E., & Vedros, R. G. Survey of Certification Methods Applied to Innovative Public School Programs. Paper presented to the Association of State Advisory Councils, Hartford, 1978.

Furguson, G. A. Statistical Analysis in Psychology and Education. (4th ed.) New York: McGraw-Hill, 1976.

Nie, N. H. SPSS—Statistical Package for the Social Sciences. (2nd ed.) New York: McGraw-Hill, 1975.

Gurel, L. The Human Side of Evaluating Human Services Programs: Problems and Prospects. In M. Guttentag & E. L. Sturening (Eds.), Handbook of Evaluation Research (Vol. 2). Beverly Hills: SAGE, 1979.

Guttentag, M., & Struening, E. L. Handbook of Evaluation Research (2 Vols.). Beverly Hills: SAGE, 1975, 1979.

Harrow, A. J. A Taxonomy of the Psychomotor Domain. New York: David McKay Company, Inc., 1972.

Horst, D. P., & Fagan, B. M. Common Evaluation Hazards. Mountainview: RMC Corporation, 1976.

Kerlinger, F. N. Foundations of Behavioral Research. New York: Holt, Rinehart & Winston, 1973.

Kerlinger, F., & Pedhazur, E. Multiple Regression in Behavioral Research. New York: Holt, Rinehart & Winston, 1973.

Krathwhol, D., & Bloom, B. Taxonomy of Educational Objectives, Handbook: Affective Domain. New York: David McKay Company, Inc., 1964.

Learned, W. S., & Hawkes, A. R. An Experiment in Responsible Learning. Boston: Merrymount Press, 1940.

Lord, F. M., & Novick, M. R. Statistical Theories of Mental Test Scores. Reading: Addison-Wesley, 1968.

Martin, P., & Pinck, R. Curriculum Improvement and Innovation: A Partnership of Students, School Teachers and Research Scholars. Cambridge: Robin & Bently, 1966.

Massachusetts Department of Education. Discretionary Grants, A Draft. Unpublished manuscript, 1979. (Available from 31 St. James St., Boston.)

Massachusetts Department of Education. Massachusetts Validation Process, An Overview. Boston: Author, 1977.

Morris, L. L., & Fitz-Gibbon, C. T. Evaluator's Handbook. Beverly Hills: SAGE, 1978.

National Diffusion Network. Educational Programs That Work. San Francisco: Far West Laboratory for Educational Research and Development, 1979.

Nunnally, J. C. Psychometric Theory. New York: McGraw-Hill, 1978.

Nunnally, J. C. The Study of Change in Evaluation Research: Principles Concerning Measurement, Experimental Design, and Analysis. In M. Guttentag, & E. L. Struening (Eds.), Handbook of Evaluation Research (Vol. 1). Beverly Hills: SAGE, 1975.

Nunnally, J. C., & Durham, R. L. Validity, Reliability, and Sepcial Problems of Measurement in Evaluation Research. In M. Guttentag & E. L. Struening (Eds.), Handbook of Evaluation Research (Vol. 1). Beverly Hills: SAGE, 1975.

Nunnally, J. C., & Wilson, W. H. Method and Theory for Developing Measures in Evaluation Research. In M. Guttentag, & E. L. Struening (Eds.), Handbook of Evaluation Research (Vol. 1). Beverly Hills: SAGE, 1975.

Patton, M. Q. Qualitative Evaluation Methods. Beverly Hills: SAGE, 1980.

Plakos, M., Plakos, J., & Babcock, R. Workbook on Program Evaluation. Princeton: Educational Testing Service, 1978.

RMC Research Corporation. ESEA Title I, Evaluation and Reporting System. Mountain View: Author, 1976.

Rand Corporation. Federal Programs Supporting Educational Change (5 Vols.). Santa Monica: Author, 1975.

Reynolds, J. F. An Historical Descriptive Assessment of the Performance Objective Program: A Systems Approach to Individualizing Instructive. Doctoral Dissertation, University of Massachusetts, 1972.

Roberts, A. O. Out-of-Level Testing. Mountain View: RMC Research Corporation, 1976.

Rogers, E. M., & Shoemaker, F. F. Communication of Innovations. New York: The Free Press, 1971.

Rossi, P. H., Freeman, H. E., & Wright, S. R. Evaluation: A Systematic Approach. Beverly Hills: SAGE, 1979.

Scirven, M. The Methodology of Evaluation. In B. R. Worthen, & J. R. Sanders (Eds.), Educational Evaluation: Theory and Practice. Belmont: Wadsworth, 1973.

Sellitz, C., Wrightsman, L. S., & Cook, S. W. Research Methods in Social Relations. New York: Holt, Rinehart & Winston, 1976.

Seroski, M. A., Lavin, R., & Burns, T. Evaluation Improvement Project. Chelmsford: Merrimack Education Center, 1979.

Sjoberg, G. Politics, Ethics and Evaluation Research. In M. Guttentab & E. L. Struening (Eds.), Handbook of Evaluation Research (Vol. 2). Beverly Hills: SAGE, 1979.

Smith, H. W. Strategies of Social Research. Englewood Cliffs: Prentice-Hall, 1975.

Smith, E. R., & Tyler, R. W. Appraising and Recording Student Progress. New York: Harper & Row, 1942.

Scriven, M. The Methodology of Evaluation. In R. Worthen & W. Sanders (Eds.), Evaluation: Theory and Practice. Belmont: Wadsworth, 1973.

Stake, R. E. The Countenance of Educational Evaluation. Teachers College Record, 1967, 68, 523-540.

Steiger, J., Find, A., & Koseecoff, J. Evaluating Federal Education Programs. Washington, DC: Capitol, 1976.

Stufflebeam, D. L. Educational Evaluation and Decision Making. Phi Delta Kappa National Study on Evaluation. Itasca: Peacock, 1971.

Swaminathan, H., Hambleton, R. K., & Cook, L. Evaluation Methods for Early Childhood Program Personnel. In Don Streets (Ed.), Administrative Handbook for Nursery Schools and Day Care Centers. Boston: Allyn and Bacon, 1979.

Tallmadge, G. K. IDEABOOK, The Joint Dissemination Review Panel. Washington: U.S. Government Print Office, 1977.

Tallmadge, G. K. Interpreting NCEs. Mountain View: RMC Research Corporation, 1976.

Tallmadge, G. K. The Regression Effect. Mountain View: RMC Research Corporation, 1976.

Tallmadge, G. K., & Horst, D. P. Statistical Adjustments for Nonequivalent Control Groups. Mountain View: RMC Research Corporation, 1976.

Thorndike, R. L. Educational Measurement. (2nd ed.) Washington, DC: American Council on Education, 1971.

Timm, N. H. Multivariate Analysis with Applications in Education and Psychology. Monterey: Brooks/Cole, 1975.

Tuckman, B. W. Evaluating Instructional Programs. Boston: Allyn & Bacon, 1979.

Tyler, R., Gagne, R., & Scriven, M. The Methodology of Evaluation. In AERA Monograph 1, Perspectives of Curriculum Evaluation. Chicago: Rand McNally, 1967.

United States General Accounting Office. Assessing Social Program Impact Evaluations: A Checklist Approach. Washington, DC: Author, 1978.

United States Office of Education. A Handbook of Identification, Validation and Dissemination of School Business Practices. Chicago: Research Corporation Association of School Business Officials, 1979.

Walker, M., & Wolf, T. Formative Evaluation. Boston: Massachusetts Department of Education, 1979.

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. Unobtrusive Measures: Nonreactive Research in the Social Sciences. Chicago: Rand McNally, 1966.

Weiss, C. H. Evaluation Research in the Political Context. In M. Guttentag & E. L. Struening (Eds.), Handbook of Evaluation Research (Vol. 1). Beverly Hills: SAGE, 1975.

Widmer, J. L. What Makes Innovation Work in Massachusetts? Boston: Massachusetts Department of Education, 1975.

Williams, M. K. Educational Manager's Guide to Project Evaluation. Andover: The NETWORK, 1978.

Wolf, T. E. Evaluating Innovative Programs. Boston: Massachusetts Department of Education, 1978.

Wolf, W. C. An Evaluation of the Massachusetts ESEA Title III Diffusion Program. Unpublished manuscript. (Available from 31 Pray Street, Amherst.)

Wolf, W. C., & Fiornio, A. J. Some Perspectives of Educational Change. The Educational Forum, November 1973.

Wood, C. T., & Tallmadge, G. K. Local Norms. Mountain View: RMC Research Corp., 1976.

Worthen, B. R., & Sanders, J. R. Educational Evaluation: Theory and Practice. Belmont: Wadsworth, 1973.

APPENDIX A

1979 Massachusetts Title IV-C

Assessment Questionnaires

PROJECT DIRECTOR QUESTIONNAIRE

1979
FORM - A

Director Name_____

Project_____

LEA_____

Region

____ Pittsfield Region
____ Springfield Region
____ Central Region
____ Southeast Region
____ Northeast Region
____ Greater Boston Region I
____ Greater Boston Region II

DIRECTIONS: The following questions have been prepared to gather data on the collective experience of Title IV-C project directors in educational innovation and management. This questionnaire is one part of the Massachusetts State Title IV Part C evaluation. Please respond to each item and return the questionnaire to the Statewide evaluation interview person assigned to your project, at the time of an on site visit. Please note that individual responses will remain confidential. Your assistance is most appreciated.

1. What is the single most important factor in the successful operation of your project?

2. What were (are) the major problems you faced while trying to make your project visible?

3. To what degree are you personally involved with the day-to-day operation of this project?

Not Involved				Very Involved
1	2	3	4	5

4. Do you make decisions concerning what groups (e.g., teachers, administrators, students) may participate in different project activities?

_____ YES _____ NO

PROJECT DIRECTOR QUESTIONNAIRE

1979
FORM - B

Director Name_____

Project_____

LEA_____

Region

____ Pittsfield Region
____ Springfield Region
____ Central Region
____ Northeast Region
____ Greater Boston Region I
____ Greater Boston Region II
____ Southeast Region

DIRECTIONS: The following questions have been prepared to gather data on the collective experience of Title IV-C project directors in educational innovation and management. This questionnaire is one part of the Massachusetts State Title IV Part C evaluation. Please respond to each item and return the questionnaire in the attached envelope. Please note that individual responses will remain confidential. Your assistance is most appreciated.

1. What is the single most important factor in the successful operation of your project?

2. What were (are) the major problems you faced while trying to make your project visible?

3. To what degree are you personally involved with the day-to-day operation of this project?

Not Involved				Very Involved
1	2	3	4	5

4. Do you make decisions concerning what groups (e.g., teachers, administrators, students) may participate in different project activities?

_____ YES _____ NO

5. What percentage of total operating costs (federal and local funds) have been budgeted for by your school district for the next fiscal year?

_____%

6. How many staff members do you currently have in your project?

Number of Staff Members	Full-Time Equivalents

7. How many of your staff members worked with your project before you received Title IV-C funds?

Number of Staff Members	Full-Time Equivalents

8. What are the three major benefits that your school system has received from your project?

1.

2.

3.

9. What misunderstandings, if any, have occured between your project and your school system?

10. Which of the following have encouraged the growth and success of your project?

	No Encouragement	Little	Neutral	Some	Much Encouragement
Superintendent	____	____	____	____	____
Other Central Office Administrators	____	____	____	____	____
School Board Members	____	____	____	____	____
School Principals	____	____	____	____	____
Teachers	____	____	____	____	____
Parents	____	____	____	____	____
Town Officials	____	____	____	____	____
State Program Officer	____	____	____	____	____
Development Center	____	____	____	____	____
Other _____	____	____	____	____	____

11. How innovative is your school district in terms of adopting new educational practices and ideas?

Not Innovative				Very Innovative
1	2	3	4	5

12. For what reasons do you usually communicate with your Regional Program Officer?

13. To what extent do you agree or disagree with the following statements?

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Not Applicable
I believe that the project will be successful	____	____	____	____	____	____
I believe that the project will be validated for Massachusetts diffusion	____	____	____	____	____	____
I believe that the project will be validated for Federal Dessemination (pass J.D.R.P.)	____	____	____	____	____	____
Project personnel believe that the project will be successful	____	____	____	____	____	____
Project personnel believe that the project will be validated for Massachusetts Diffusion	____	____	____	____	____	____
Project personnel believe that the project will be validated for Federal Dessemination (pass J.D.R.P.)	____	____	____	____	____	____
This project runs smoothly within my school	____	____	____	____	____	____
This project is fulfilling a major need	____	____	____	____	____	____
This project has aroused controversy	____	____	____	____	____	____
Project implementation has created some major problems	____	____	____	____	____	____

14. Have your project objectives been modified since your project began?

_____ YES _____ NO

15. How would you rate the evaluation component of your project?

Excellent Very Good Good Poor Very Poor

16. What areas of your project are being evaluated?

_____ Student achievement

_____ Student attitude

_____ Other student changes (e.g., social, physical)

_____ Student skill development

_____ Personnel/Teacher achievement

_____ Personnel/Teacher attitude

_____ Personnel/Teacher skill development

_____ Cost effectiveness

_____ Parent attitude

_____ Community attitude

_____ Other_____

17. The following describes my feelings about the Massachusetts Validation process :

_____ Helpful _____ Unnecessary

_____ Hinderance _____ An Opportunity

_____ Necessary _____ No Feeling

18. For each of the people, or groups of people, listed below, indicate how frequently you have personally talked with each.

		FREQUENCY OF CONTACT			
		Daily	Weekly	Monthly	Less Frequent
A.	Superintendent	____	____	____	____
B.	Other central Office administrators	____	____	____	____
C.	School Board Members	____	____	____	____
D.	School Principals	____	____	____	____
E.	School Specialists	____	____	____	____
F.	Non-project teachers	____	____	____	____
G.	Project teachers	____	____	____	____
H.	Participating students	____	____	____	____
I.	Parents of participating students	____	____	____	____
J.	Other Parents	____	____	____	____
K.	Town Officials	____	____	____	____
L.	Public Media People	____	____	____	____
M.	State program officer	____	____	____	____
N.	Project evaluator	____	____	____	____
O.	HEC Title IV Development Center Staff	____	____	____	____
P.	Other Consultants	____	____	____	____
Q.	Other_____	____	____	____	____

19. For each of the people or groups of people listed below, indicate whether you are satisfied with the amount of personal contact you presently have with each.

	SATISFACTION WITH CONTACT		
	Satisfied	Want More Contact	Want Less Contact
A. Superintendents	____	____	____
B. Other central office administrators	____	____	____
C. School Board Members	____	____	____
D. School Principals	____	____	____
E. School Specialists	____	____	____
F. Non-project teachers	____	____	____
G. Project teachers	____	____	____
H. Participating students	____	____	____
I. Parents of participating students	____	____	____
J. Other Parents	____	____	____
K. Town Officials	____	____	____
L. Public Media People	____	____	____
M. State Program Officer	____	____	____
N. Project Evaluator	____	____	____
O. HEC Title IV Development Center Staff	____	____	____
P. Other Consultants	____	____	____
Q. Other_____	____	____	____

20. For each of the individuals or groups listed below, indicate whether you are satisfied with their level of understanding of your project.

	SATISFACTION WITH UNDERSTANDING		
	Satisfied With Understanding	Dissatisfied With Understanding	N/A
A. Superintendent	____	____	____
B. Other central office administrators	____	____	____
C. School Board Members	____	____	____
D. School Principals	____	____	____
E. School Specialists	____	____	____
F. Non-project teachers	____	____	____
G. Project teachers	____	____	____
H. Participating students	____	____	____
I. Parents of participating students	____	____	____
J. Other Parents	____	____	____
K. Town Officials	____	____	____
L. Public Media People	____	____	____
M. State Program Officer	____	____	____
N. Project Evaluator	____	____	____
O. HEC Title IV Development Center Staff	____	____	____
P. Other Consultants	____	____	____
Q. Other_____	____	____	____

21. How helpful has your Regional Program Officer been with:

	Not At All Helpful	Only Slightly Helpful	Neutral	Heloful	Very Helpful	Not Aoplicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Preparation For Validation	____	____	____	____	____	____
F. Dealing with required procedures	____	____	____	____	____	____

22. How helpful have your school system personnel been with:

	Not At All Helpful	Only Slightly Heloful	Neutral	Heloful	Very Heloful	Not Aoplicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Preparation For Validation	____	____	____	____	____	____
F. Dealing with required procedures	____	____	____	____	____	____

23. How helpful has technical assistance from the HEC Title IV-C Development Center been with:

	Not At All Helpful	Only Slightly Helpful	Neutral	Helpful	Very Helpful	Not Applicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Preparation For Validation	____	____	____	____	____	____
F. Dealing with required procedures	____	____	____	____	____	____

24. How helpful has your project evaluator been with:

	Not At All Helpful	Only Slightly Helpful	Neutral	Helpful	Very Helpful	Not Applicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Preparation For Validation	____	____	____	____	____	____
F. Dealing with required procedures	____	____	____	____	____	____

25. In what areas do you feel you need training?

	No Need For Training	Need Training	Received Training
A. Change Theory (e.g., principals, strategies)	_____	_____	_____
B. Leadership Theory (e.g., techniques, styles)	_____	_____	_____
C. Internal Communication (e.g. staff development, group process)	_____	_____	_____
D. Resource Utilization (e.g. time, personnel)	_____	_____	_____
E. Financial Management (e.g. budget, bookkeeping)	_____	_____	_____
F. Formative Evaluation (e.g. purpose, type, implementation)	_____	_____	_____
G. Summative Evaluation (e.g. purpose, type)	_____	_____	_____
H. Data Utilization (e.g. analysis, reporting)	_____	_____	_____
I. Massachusetts Validation Requirements	_____	_____	_____
J. Federal Validation Requirements	_____	_____	_____
K. Marketing Techniques	_____	_____	_____
L. Packaging	_____	_____	_____
M. Diffusion Services (e.g. diffusing project to another site)	_____	_____	_____
N. Other (Please Specify) _____	_____	_____	_____

PROJECT DIRECTOR QUESTIONNAIRE

1979
FORM - C

Director Name _____

Project _____

LEA _____

Region

____ Pittsfield Region
____ Springfield Region
____ Central Region
____ Northeast Region
____ Greater Boston Region I
____ Greater Boston Region II
____ Southeast Region

DIRECTIONS: The following questions have been prepared to gather data on the collective experience of Title IV-C project directors in educational innovation and management. This questionnaire is one part of the Massachusetts State Title IV Part C evaluation. Please respond to each item and return the questionnaire to the Statewide evaluation interview person assigned to your project, at the time of an on site visit. Please note that individual responses will remain confidential. Your assistance is most appreciated.

1. What is the single most important factor in the successful operation of your project?

PROJECT DIRECTOR QUESTIONNAIRE

1979
FORM - D

Director Name_____

Project_____

LEA_____

Region

____ Pittsfield Region
____ Springfield Region
____ Central Region
____ Northeast Region
____ Greater Boston Region I
____ Greater Boston Region II
____ Southeast Region

DIRECTIONS: The following questions have been prepared to gather data on the collective experience of Title IV-C project directors in educational innovation and management. This questionnaire is one part of the Massachusetts State Title IV Part C evaluation. Please respond to each item and return the questionnaire in the attached envelope. Please note that individual responses will remain confidential. Your assistance is most appreciated.

1. What is the single most important factor in the successful operation of your project?

2. What were the major problems you faced while implementing your project?

3. To what degree are you personally involved with the day-to-day operations of this project?

Not Involved				Very Involved
1	2	3	4	5

4. To what degree will you personally be involved with the day-to-day operation of this project next year?

Not Involved				Very Involved
1	2	3	4	5

5. How many staff members do you currently have in your project?

_____ Number of Staff Members _____ Full-Time Equivalents

6. How many of your staff members worked with your project before you received Title IV-C funds?

_____ Number of Staff Members _____ Full-Time Equivalents

7. What misunderstandings, if any, have occured between your project and your school system?

8. What are the three major benefits that your school system has received from your project?

1.

2.

3.

9. Which of the following have encouraged the growth and success of your project?

	No Encouragement	Little	Neutral	Some	Much Encouragement
Superintendent	____	____	____	____	____
Other Central Office Administrators	____	____	____	____	____
School Board Members	____	____	____	____	____
School Principals	____	____	____	____	____
Teachers	____	____	____	____	____
Parents	____	____	____	____	____
Town Officials	____	____	____	____	____
State Program Officer	____	____	____	____	____
Development Center	____	____	____	____	____
Other_____	____	____	____	____	____

10. Do you make decisions concerning what groups (e.g., teachers, administrators, students) may participate in different project activities?

_____YES _____NO

11. What percentage of total operating costs (federal and local funds) have been budgeted for by your school district for the next fiscal year?

_____%

12. To what extent do you agree or disagree with the following statements?

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	N/A
This project runs smoothly within my school	____	____	____	____	____	____
This project is fulfilling a major need	____	____	____	____	____	____
This project has arroused controversy	____	____	____	____	____	____
Project implementation has created some major problems	____	____	____	____	____	____

13. How innovative is your school district in terms of adopting new educational practices and ideas?

Not Innovative				Very Innovative
1	2	3	4	5

14. Is there a reward mechanism in place in your system which encourages new ideas?

_____ YES _____ NO

15. What is the liklihood that your project will continue in your system next year?

Positively	Very Likely	Maybe	Unlikely	No Chance

16. Are the teachers in your system usually involved in decisions to try out new ideas?

_____ YES _____ NO

17. When Title IV-C funds stop will others in your system still use the ideas, methods and materials of your project?

_____ YES _____ NO

18. What materials will you use to diffuse your project?

19. Are project materials packaged and ready for diffusion?

20. Have you identified the sites you will be diffusing to?

_____ YES _____ NO

21. How would you rate your projects readiness for diffusion?

Not Ready For Diffusion				Ready For Diffusion
1	2	3	4	5

22. How many sites will you be diffusing to?

_____

23. What methods will you be using to diffuse to other sites?

24. Have your project objectives been modified for diffusion?

_____ YES _____ NO

25. Is your project being evaluated for presentation to the Joint Dissemination Review Panel?

_____ YES _____ NO

26. For what reasons do you usually communicate with your Regional Program Officer?

27. How helpful has technical assistance from the HEC Title IV-C Development Center been with:

	Not At All Helpful	Only Slightly Helpful	Neutral	Helpful	Very Helpful	Not Applicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Dealing with Required Procedures	____	____	____	____	____	____
F. Preparing for Project Diffusion	____	____	____	____	____	____
G. Identifying Dissusion Sites	____	____	____	____	____	____

28. How helpful has your Regional Program Officer been with:

	Not At All Helpful	Only Slightly Helpful	Neutral	Helpful	Very Helpful	Not Applicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Preparation For Validation	____	____	____	____	____	____
F. Dealing with Required Procedures	____	____	____	____	____	____
G. Preparing for Project Diffusion	____	____	____	____	____	____
H. Identifying Diffusion Sites	____	____	____	____	____	____

29. How helpful have your in school personnel been with:

	Not At All Helpful	Only Slightly Helpful	Neutral	Helpful	Very Helpful	Not Applicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Preparation for Validation	____	____	____	____	____	____
F. Dealing with Required Procedures	____	____	____	____	____	____
G. Preparing for Project Diffusion	____	____	____	____	____	____
H. Identifying Diffusion Sites	____	____	____	____	____	____

30. How helpful has your project evaluator been with:

	Not At All Helpful	Only Slightly Helpful	Neutral	Helpful	Very Helpful	Not Applicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Preparation for Validation	____	____	____	____	____	____
F. Dealing with Required Procedures	____	____	____	____	____	____
G. Preparing for Project Diffusion	____	____	____	____	____	____
H. Identifying Diffusion Sites	____	____	____	____	____	____

31. In which of the following areas do you feel you need training?

	No Need For Training	Need Training	Received Training
A. Change Theory (e.g. principals, strategies)	_____	_____	_____
B. Leadership Theory (e.g. techniques, Styles)	_____	_____	_____
C. Internal Communication (e.g. staff development, group process)	_____	_____	_____
D. Resource Utilization (e.g. time personnel)	_____	_____	_____
E. Financial Management (e.g. budgets, bookkeeping)	_____	_____	_____
F. Formative Evaluation (e.g. purpose, types implementation)	_____	_____	_____
G. Summative Evaluation (e.g. purpose, type)	_____	_____	_____
H. Data Utilization (e.g. analysis, reporting)	_____	_____	_____
I. Massachusetts Validation Requirements	_____	_____	_____
J. Federal Validation Requirements	_____	_____	_____
K. Marketing Techniques	_____	_____	_____
L. Packaging	_____	_____	_____
M. Diffusion Services (e.g. diffusing project to another site)	_____	_____	_____
N. Other (Please Specify) _____	_____	_____	_____

PROJECT DIRECTOR QUESTIONNAIRE

1979
FORM - E

Director Name_____

Project_____

LEA_____

Region

____ Pittsfield Region
____ Springfield Region
____ Central Region
____ Northeast Region
____ Greater Boston Region I
____ Greater Boston Region II
____ Southeast Region

DIRECTIONS: The following questions have been prepared to gather data on the collective experience of Title IV - C project directors in educational innovation and management. This questionnaire is one part of the Massachusetts State Title IV Part C evaluation. Please respond to each item and return the questionnaire to the Statewide evaluation interview person assigned to your project, at the time of an on site visit. Please note that individual responses will remain confidential. Your assistance is most appreciated.

1. What is the single most important factor in the successful operation of your project?

2. What were the major problems you faced while implementing your project?

3. To what degree are you personally involved with the day-to-day operation of this project?

Not Involved				Very Involved
1	2	3	4	5

PROJECT DIRECTOR QUESTIONNAIRE

1979
FORM - F

Director Name_____

Project_____

LEA_____

Region

____ Pittsfield Region
____ Springfield Region
____ Central Region
____ Northeast Region
____ Greater Boston Region I
____ Greater Boston Region II
____ Southeast Region

DIRECTIONS: The following questions have been prepared to gather data on the collective experience of Title IV-C project directors in educational innovation and management. This questionnaire is one part of the Massachusetts State Title IV Part C evaluation. Please respond to each item and return the questionnaire in the attached envelope. Please note that individual responses will remain confidential. Your assistance is most appreciated.

1. What is the single most important factor in the successful operation of your project?

2. What were the major problems you faced while implementing your project?

3. To what degree are you personally involved with the day-to-day operation of this project?

Not Involved				Very Involved
1	2	3	4	5

4. To what degree will you personally be involved with the day-to-day operation of this project next year?

Not Involved				Very Involved
1	2	3	4	5

5. How many staff members do you currently have in your project?

_____ _____
Number of Staff Members Full-Time Equivalents

6. How many of your staff members worked with your project before you received Title IV - C funds?

_____ _____
Number of Staff Members Full-Time Equivalents

7. What are the three major benefits that your school system has received from your project?

1.

2.

3.

8. What misunderstandings, if any, have occurred between your project and your school system?

9. Which of the following have encouraged the growth and success of your project?

	No Encouragement	Little	Neutral	Some	Much Encouragement
Superintendent	____	____	____	____	____
Other Central Office Administrators	____	____	____	____	____
School Board Members	____	____	____	____	____
School Principals	____	____	____	____	____
Teachers	____	____	____	____	____
Parents	____	____	____	____	____
Town Officials	____	____	____	____	____
State Program Officer	____	____	____	____	____
Development Center	____	____	____	____	____
Other_____	____	____	____	____	____

10. Do <u>you</u> make decisions concerning what groups (e.g., teachers, administrators, students) may participate in different project activities?

_____ YES _____ NO

11. What percentage of total operating costs (federal and local funds) have been budgeted for by your school district for the next fiscal year?

_____ %

12. To what extent do you agree or disagree with the following statements?

	Strongly Agree	Agree	Neutral	Disagree	Strongly Disagree	Not Applicable
This project runs smoothly within my school	____	____	____	____	____	____
This project is fulfilling a major need	____	____	____	____	____	____
This project has aroused controversy	____	____	____	____	____	____
Project implementa has created some major problems	____	____	____	____	____	____

13. How innovative is your school district in terms of adopting new educational practices and ideas?

Not Innovative				Very Innovative
1	2	3	4	5

14. Is there a reward mechanism in place in your system which encourages new ideas?

_____ YES _____ NO

15. Are the teachers in your system usually involved in decisions to try out new ideas?

_____ YES _____ NO

16. What is the likelihood that your project will continue in your system next year?

Positively	Very Likely	Maybe	Unlikely	No Chance

17. When Title IV-C funds stop will others in your system still use the ideas, methods and materials of your project?

_____ YES _____ NO

18. What problems are you facing as Federal Funding ends?

19. Is your system planning to apply for Title IV-C funds in the future?

_____ YES _____ NO _____ DON'T KNOW

20. For what reasons do you usually communicate with your Regional Program Officer?

21. How helpful has technical assistance from the HEC Title IV-C Development Center been with:

	Not At All Helpful	Slightly Helpful	Neutral	Helpful	Very Helpful	Not Applicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Preparation For Validation	____	____	____	____	____	____
F. Dealing With Required Procedures	____	____	____	____	____	____

22. How helpful has your Regional Program Officer been with:

	Not At All Helpful	Slightly Helpful	Neutral	Helpful	Very Helpful	Not Applicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Preparation For Validation	____	____	____	____	____	____
F. Dealing With Required Procedures	____	____	____	____	____	____

23. How helpful have your in school personnel been with:

	Not At All Helpful	Slightly Helpful	Neutral	Helpful	Very Helpful	Not Applicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Preparation For Validation	____	____	____	____	____	____
F. Dealing With Required Procedures	____	____	____	____	____	____

24. How helpful has your project evaluator been with:

	Not At All Helpful	Slightly Helpful	Neutral	Helpful	Very Helpful	Not Applicable
A. Program Planning	____	____	____	____	____	____
B. Program Implementation	____	____	____	____	____	____
C. Program Evaluation	____	____	____	____	____	____
D. Training/Skill Development	____	____	____	____	____	____
E. Preparation For Validation	____	____	____	____	____	____
F. Dealing With Required Procedures	____	____	____	____	____	____

1978 Massachusetts Title IV-C
Assessment Questionnaire

Project Director Questionnaire

Director Name_____

Project_____

Region_____

Directions: The following questions have been prepared to gather data on the collective experience of Title IV-c project directors in educational innovation and management. This questionnaire is one part of the Massachusetts State Title IV Part C evaluation. Please respond to each item and return the questionnaire to the Statewide evaluation interview person assigned to your project. Please note that individual responses will remain confidential. Your assistance is most appreciated.

Background:

1. How did you come to be director of your project?

___ helped plan project

___ wrote proposal

___ hired from within system

___ hired from outside the system

___ other_____

2. What was your position immediately prior to your appointment as director of this Title IV-c project? Please check appropriate response.

___ Teacher

___ Department Chairperson

___ Guidance Counselor

___ School Building Administrator

___ School District Administrator

___ Other (please specify:_____)

3. As of June 1978, how many years will you have been in each of the following?

	less than 1	1-5	5-10	over 10
Title IV-c project				
In this district				
Total in education profession				

4. What is the highest degree you have earned? (Please specify field)

___ Bachelors Degree_____

___ Masters Degree_____

___ Certificate of Advanced Graduate Study_____

___ Doctoral Degree_____

5. Have you had other professional experience with programs similar in content to your present project such as:

	Yes	No
A. project director	___	___
B. staff member	___	___
C. evaluator	___	___
D. administrator	___	___
E. other_____	___	___

6. How many years of experience have you had in other professional experiences with innovative educational projects such as:

	Years
A. project director	____
B. staff member	____
C. evaluator	____
D. administrator	____
E. other_____	____

7. What is the extent of your current knowledge and/or practical understanding of each of the following management-related topics?

	None				Extensive
A. Change theory (e.g. principles, strategies)	1	2	3	4	5
B. Leadership theory (e.g. techniques, styles)	1	2	3	4	5
C. Internal communication (e.g. staff development, group process)	1	2	3	4	5
D. External communication (e.g. dissemination, diffusion)	1	2	3	4	5
E. Resource utilization (e.g. time, budget, personnel)	1	2	3	4	5
F. Evaluation theory (e.g. purposes, types)	1	2	3	4	5
G. Evaluation application (e.g. instrument design, implementation)	1	2	3	4	5
H. Data utilization (e.g. analysis, reporting)	1	2	3	4	5
I. Other (please specify) _____	1	2	3	4	5

8. On the five point scale, how would you rate your ease in dealing with each of the following areas?

	Difficulty in Handling			No Difficulty		N/A
A. obtaining initial funding	1	2	3	4	5	___
B. obtaining project approval	1	2	3	4	5	___
C. recruiting appropriate staff	1	2	3	4	5	___
D. scheduling project staff	1	2	3	4	5	___
E. scheduling students	1	2	3	4	5	___
F. recruiting participants	1	2	3	4	5	___
G. retaining participants	1	2	3	4	5	___
H. obtaining appropriate facilities	1	2	3	4	5	___

8. Continued.....

	Difficulty in Handling				No Difficulty	N/A
I. obtaining relevant materials	1	2	3	4	5	___
J. scheduling training workshops	1	2	3	4	5	___
K. scheduling parent meetings	1	2	3	4	5	___
L. obtaining continued funding	1	2	3	4	5	___
M. obtaining organizational support	1	2	3	4	5	___
N. obtaining community support	1	2	3	4	5	___

9. How do you feel your position, project director, will be preserved at the conclusion of the Federal funding phase?

___ full time Director

___ part time Director

___ included in the responsibility of another administrator

___ included in the responsibility of a teacher position

___ uncertain

___ no position

10. What percentage of the total operating costs (Federal and local funds) have been budgeted for by your school district for the next fiscal year?

_____ %

Dissemination:

11. What means do you use to inform others of your project? (select all appropriate responses)

___ personal contact

___ T.V.

___ Radio

___ Press

___ Brochures

___ Newsletters

___ Other_____

12. For each of the people, or groups of people, listed below, indicate how frequently you have personally talked with each.

	Frequency of Contact			
	daily	weekly	monthly	less frequent
A. Superintendent	___	___	___	___
B. Other central office administrators	___	___	___	___
C. School board members	___	___	___	___
D. School principals	___	___	___	___
E. School specialists	___	___	___	___
F. Non-project teachers	___	___	___	___
G. Project teachers	___	___	___	___
H. Participating students	___	___	___	___
I. Parents of participating students	___	___	___	___
J. Other parents	___	___	___	___
K. Town officials	___	___	___	___
L. Public media people	___	___	___	___
M. State program officer	___	___	___	___
N. Project evaluator	___	___	___	___
O. HEC Title IV Development Center staff	___	___	___	___
P. Other consultants	___	___	___	___
Q. Other_____	___	___	___	___

13. For each of the people or groups of people listed below, indicate whether you are satisfied with the amount of personal contact you presently have with each.

	Satisfaction with Contact		
	satisfied	want more contact	want less contact
A. Superintendent	____	____	____
B. Other central office administrators	____	____	____
C. School board members	____	____	____
D. School principals	____	____	____
E. School specialists	____	____	____
F. Non-project teachers	____	____	____
G. Project teachers	____	____	____
H. Participating students	____	____	____
I. Parents of participating students	____	____	____
J. Other parents	____	____	____
K. Town officials	____	____	____
L. Public media people	____	____	____
M. State program officer	____	____	____
N. Project evaluator	____	____	____
O. HEC Title IV Development Center staff	____	____	____
P. Other consultants	____	____	____
Q. Other_____	____	____	____

Project Objectives:

14. Have your project objectives been modified since your project began?

____Yes ____No

15. For each of the individuals or groups listed below, indicate whether you are satisfied with their level of understanding of your project.

	Satisfaction with Understanding		
	Satisfied with Understanding	Dissatisfied with Understanding	N/A
A. Superintendent	___	___	___
B. Other central office administrators	___	___	___
C. School board members	___	___	___
D. School principals	___	___	___
E. School specialists	___	___	___
F. Non-project teachers	___	___	___
G. Project teachers	___	___	___
H. Participating students	___	___	___
I. Parents of participating students	___	___	___
J. Other parents	___	___	___
K. Town officials	___	___	___
L. Public media people	___	___	___
M. State program officer	___	___	___
N. Project evaluator	___	___	___
O. HEC Title IV Development Center staff	___	___	___
P. Other consultants	___	___	___
Q. Other_____	___	___	___

Project Success:

16. What is the single most important factor in the successful operation of your project?

17. From the perspective of your project, how would you define success?

18. To what degree of success, as defined by your own terms, do you anticipate your project will attain by the end of three years?

very high | high | medium | low | very low

Evaluation:

19. How would you rate the evaluation component of your project at this time?

excellent | very good | good | poor | very poor

20. The following describes my feelings about the Massachusetts Validation Process

___ helpful
___ hinderance
___ necessary
___ unnecessary
___ an opportunity
___ no feeling

Technical Assistance:

21. To what degree do you feel the HEC Title IV-c Development Center has provided assistance with:

	very low	low	medium	high	very high
A. Proposal writing	___	___	___	___	___
B. Program planning	___	___	___	___	___
C. Program implementation	___	___	___	___	___
D. Program evaluation	___	___	___	___	___
E. Training/skill development	___	___	___	___	___
F. Preparation for validation	___	___	___	___	___
G. Dealing with required procedures	___	___	___	___	___

22. To what degree do you feel your Regional Program Officer has provided you assistance with:

	very low	low	medium	high	very high
A. Proposal writing	___	___	___	___	___
B. Program planning	___	___	___	___	___
C. Program implementation	___	___	___	___	___
D. Program evaluation	___	___	___	___	___
E. Training/skill development	___	___	___	___	___
F. Preparation for validation	___	___	___	___	___
G. Dealing with required procedures	___	___	___	___	___

23. To what degree do you feel your in-school personnel has provided assistance with:

	very low	low	medium	high	very high
A. Proposal writing	___	___	___	___	___
B. Program planning	___	___	___	___	___
C. Program implementation	___	___	___	___	___
D. Program evaluation	___	___	___	___	___
E. Training/skill development	___	___	___	___	___
F. Preparation for validation	___	___	___	___	___
G. Dealing with required procedures	___	___	___	___	___

24. To what degree do you feel your Project Evaluator has provided assistance with:

	very low	low	medium	high	very high
A. Proposal writing	___	___	___	___	___
B. Program planning	___	___	___	___	___
C. Program implementation	___	___	___	___	___
D. Program evaluation	___	___	___	___	___
E. Training/skill development	___	___	___	___	___
F. Preparation for validation	___	___	___	___	___
G. Dealing with required procedures	___	___	___	___	___

25. To what degree do you feel other consultants have provided assistance with:

	very low	low	medium	high	very high
A. Proposal writing	____	____	____	____	____
B. Program planning	____	____	____	____	____
C. Program implementation	____	____	____	____	____
D. Program evaluation	____	____	____	____	____
E. Training/skill development	____	____	____	____	____
F. Preparation for validation	____	____	____	____	____
G. Dealing with required procedures	____	____	____	____	____

Comments:

APPENDIX B

Validator Questionnaire

VALIDATOR QUESTIONNAIRE

Your Name_____

Project Name_____

This project was:

_____ validated for diffusion

_____ validated for dissemination

_____ not validated

1. Did the evaluation design for this project yield data that demonstrate change (growth) in the client group?

_____ Yes _____ No _____ Not Sure

Comments

2. If you answered "Yes" to #1, did the evaluation design for this project yield data that demonstrate that the project is responsible for the change?

_____ Yes _____ No _____ Not Sure

Comments

3. The data obtained for the evaluation of this project were obtained from:

_____ Nationally normed instruments

_____ Locally developed instruments

_____ Both

4.A. If locally developed instruments were used to collect data, were the procedures used to develop the instruments explained?

_____ Yes _____ No

B. Are the instruments reliable? ___Yes ___No ___Not Sure ___Not Applicable

C. Are the instruments valid? ___Yes ___No ___Not Sure ___Not Applicable

5. If statistical procedures were used to analyze/compare data, were they properly employed? ___Yes ___No ___Not Sure

6. Was the "Validation Team Report" a useful process?

_____ Yes _____ No _____ Not Sure

7. What were the strengths of the "Validation Team Report" document?

8. What were the weaknesses of the "Validation Team Report" document?

Please return to:

Mary Seroski, Evaluator
Title IV Development Center
321 Main Street
Amherst, MA 01002

APPENDIX C

Regional Program Officer

On-Site Questionnaire

REGIONAL PROGRAM OFFICER

ON-SITE QUESTIONNAIRE

Region _____

Regional Program Officer _____

1. What are the strengths and weaknesses of the Title IV-C staff plan for monitoring funded projects?

Strengths:

Weaknesses:

Recommendations:

2. What are the strengths and weaknesses of the Title IV-C staff plan for evaluating funded projects?

Strengths:

Weaknesses:

Recommendations:

3. What are the strengths and weaknesses of the Title IV-C staff plan for disseminating successful (validated for diffusion) projects?

Strengths:

Weaknesses:

Recommendations:

4. What are the strengths and weaknesses of the Title IV-C staff plan for providing technical assistance to the public and non-public schools in your region?

Strengths:

Weaknesses:

Recommendations:

5. What is your greatest frustration as a regional program officer?

6. What are some recommendations you would make to strengthen the implementation of Title IV-C in the Commonwealth?

APPENDIX D

Evaluation Resource Center

Amherst, Massachusetts

Evaluation Resource Center
Amherst, Massachusetts

PROJECT SCALE

PROJECT NAME _____

PROJECT COMMUNITY/COLLABORATIVE _____

DIRECTOR NAME _____

R.P.O. NAME _____

DATE _____

Following are nine statements. Please place a check in the column labeled "True" if you consider the statement a valid description of your project at this time, if not, place a check in the column labeled "False".

True	False	
_____	_____	1. Measurable student learning objectives for this project have been established.
_____	_____	2. The target population has been identified.
		a. If True, is your population over a three year period... ____single? ____multiple?
_____	_____	3. The project director is knowledgeable about project needs and the decision-making dynamics of the community.
_____	_____	4. The project staff has had experience with this project.
_____	_____	5. This project is widely known within its system.
_____	_____	6. The system is supportive of this project.
_____	_____	7. This project curriculum has been thoroughly developed and is currently in use.
_____	_____	8. This project has produced materials that are currently in use.
_____	_____	9. Instrumentation for the evaluation of this project is known and acceptable.

This project is in the _____ Stage.

EVALUATION RESOURCE CENTER
Amherst, Massachusetts

KEY

If you checked "True" nine times your project is in the OUTCOMES Stage.

If you checked "True" for numbers 1,2,7 and one other, or 2,3,7 and one other your project is in the IMPLEMENTATION Stage.

If you checked "True" in any other combination your project is in the INITIATION Stage.

If you checked "single" for number 2a. your project is a SINGLE POPULATION project. In this case, please contact the Evaluation Resource Center.

EVALUATION RESOURCE CENTER
Amherst, Massachusetts

Project Scale Outline

OUTCOMES

- Measurable student learning objectives have been established.
- Project is known in system.
- Project director is knowledgeable of the project needs and of the decision making dynamics of the community.
- Target population identified.
- System is supportive of project.
- Project Curriculum has been thoroughly developed and is currently in use.
- Project materials (products) are in existence.
- Project staff is experienced with project.
- Instrumentation for evaluation/validation are known and acceptable.

IMPLEMENTATION

- Project director is experienced with project.
- Project goals are NOT clear.
- Project Curriculum has been thoroughly developed and is currently in use.
- Target population identified.

OR

- Project director is NOT experienced with project.
- Project goals and/or objectives are clear.
- Project Curriculum has been thoroughly developed and is currently in use.
- Target population identified.

OR

- Any similar combination.

INITIATION

- A project that is not yet in the IMPLEMENTATION Stage.

SINGLE POPULATION

- A project that will have a single target population over a three year period.

- Immediate Summative Evaluation assistance will be needed. The project should work with Title IV-C Coordinator, Title IV-C Regional Program Officers, and the Evaluation Resource Center in the selection of the project evaluator.

E.S.E.A. TITLE IV-C

SUMMATIVE EVALUATION MANAGEMENT SYSTEM

	R.P.O.	E.R.C.	M.S.C.
1. Student goals defined.	_____		
2. Major objectives, in terms of student results, defined.	_____		
3. Existing standardized tests reviewed.	_____	_____	
4. Psychometrician may be sought.	_____		_____
5. Psychometrician selected and may be contracted. Submit: Resume, Samples of previous work, Proposal for evaluation contract, Contract	_____	_____	_____
6. Instruments are acceptable. Submit: Instruments, Description of development process, Demonstration of reliability and validity.	_____	_____	_____
7. Summative evaluator may be sought.	_____		_____
8. Evaluator selected and may be contracted. Submit: Resume, Samples of previous work, Proposal for evaluation contract including symbolic evaluation design, Contract.	_____	_____	_____
9. Interim reports of progress. Dates:	_____		
10. Data reviewed for decision on validation request. Decision jointly reached.	_____		
11. Validation request submitted.	_____		
12. Validation Team Report Received.			_____